

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:15:31 ; Search time 186 Seconds  
(without alignments)  
1322.861 Million cell updates/sec

Title: US-09-853-880A-17  
Perfect score: 3026  
Sequence: 1 MECLYFLGFLLLAARLPD.....PPGNQEKDPLLNQEFKGV 560

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*

- 1: Genesep1980s.\*
- 2: Genesep1990s.\*
- 3: Genesep2000s.\*
- 4: Genesep2001s.\*
- 5: Genesep2002s.\*
- 6: Genesep2003as.\*
- 7: Genesep2003bs.\*
- 8: Genesep2004s.\*
- 9: Genesep2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	2 AAW35382	AAW35382 Murine me
2	3026	100.0	560	3 AAB11329	Aab11329 Human lun
3	3026	100.0	560	5 ABB78200	Abb78200 Amino aci
4	3026	100.0	560	5 ABB74961	Abb74961 Human lun
5	3026	100.0	560	5 ABP61881	Abp61881 Human lun
6	3026	100.0	560	6 ABUS6592	Abus6592 Lung canc
7	3026	100.0	560	6 ABG72962	Abg72962 Human oet
8	3026	100.0	560	6 ABU70852	Abu70852 Human aci
9	3026	100.0	560	7 ADA28315	Ada28315 Human lun
10	3026	100.0	560	7 ADK95620	Adk95620 Human NOV
11	3026	100.0	560	7 ADH36879	Adh36879 Human lun
12	3026	100.0	560	7 ADJ68660	Adj68660 Human hea
13	3026	100.0	560	7 ADL14995	Adl14995 Human NNB
14	3026	100.0	560	7 ADN39940	Adn39940 Cancer/an
15	3026	100.0	560	8 ADH56342	Adh56342 Human nmb
16	3026	100.0	560	8 ADJ75569	Adj75569 Marker ge
17	3026	100.0	560	8 ADM56882	Adm56882 Human lun
18	3026	100.0	560	8 ADQ18310	Adq18310 Human sof
19	3026	100.0	560	8 ADP23127	Adp23127 PRO polyyp
20	3026	100.0	560	9 ADU98374	Adu98374 Lung tumo
21	3026	100.0	560	9 AEB10182	Aeb10182 Cancer re
22	3026	100.0	561	8 AEU06546	Aeu06546 Novel bro
23	3026	100.0	561	8 ADR66659	Adr66659 Human pro
24	3026	100.0	563	8 ADR66317	Adr66317 Human pro

25	3026	100.0	563	8 ADR66339	Adr66339 Human pro
26	3026	100.0	563	8 ADR66681	Adr66681 Human pro
27	3007	99.4	572	7 ADD78235	Add78235 Human CGD
28	3006	99.3	572	5 AAU83612	Aau83612 Human PRO
29	3006	99.3	572	5 ADY31782	Ady31782 Novel hum
30	3006	99.3	572	6 ABU80759	Abu80759 Human PRO
31	3006	99.3	572	6 ABUS33725	Abus33725 Novel hum
32	3006	99.3	572	6 ABUS2068	Abus2068 Novel hum
33	3006	99.3	572	6 ABJ72248	Abj72248 Human PRO
34	3006	99.3	572	6 ABJ72376	Abj72376 Human PRO
35	3006	99.3	572	6 ABO34271	Abo34271 Human mem
36	3006	99.3	572	7 ABJ72078	Abj72078 Human mem
37	3006	99.3	572	7 ADB83532	Adb83532 Novel hum
38	3006	99.3	572	7 ADB80638	Adb80638 Novel hum
39	3006	99.3	572	7 ADB73179	Adb73179 Novel hum
40	3006	99.3	572	7 ADB78261	Adb78261 Novel hum
41	3006	99.3	572	7 ADB84909	Adb84909 Human PRO
42	3006	99.3	572	7 ADB78015	Adb78015 Novel hum
43	3006	99.3	572	7 ADB87081	Adb87081 Human PRO
44	3006	99.3	572	7 ADB84663	Adb84663 Human PRO
45	3006	99.3	572	7 ADB83778	Adb83778 Novel hum

## ALIGNMENTS

## RESULT 1

AAW35382	ID	AAW35382 standard; protein; 560 AA.
XX	AC	AAW35382;
XX	DT	26-FEB-1998 (first entry)
XX	DE	Murine metastatic nucleic acid sequence product.
XX	DE	Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis;
KW	KW	treatment; metastasis; hyperplasia; dysplasia; hypertrophy; screening.
XX	OS	Mus musculus.
XX	PN	WO9718454-A2. ✓
XX	PD	22-MAY-1997.
XX	PF	15-NOV-1996; 96WO-US018567.
XX	PR	16-NOV-1995; 95US-0006838P.
PR	PR	30-JAN-1996; 96US-00594031.
XX	XX	(THOM/) THOMPSON T.
XX	XX	Thompson T;
XX	XX	WPI; 1997-289397/26.
DR	DR	Identifying tumour metastatic sequences - by introducing transfected
PT	PT	cells into host mammal and analysing primary and metastatic sequences by
PT	PT	differential display PCR.
XX	XX	Disclosure; Fig 12CI; 102pp; English.
PS	PS	Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse
XX	XX	embryos. The UGS cells were infected with retroviruses, cultured and
CC	CC	implanted under the renal capsule of mice. Reconstructions were harvested
CC	CC	5 weeks later, when they showed signs of distress from the tumour burden.
CC	CC	Metastatised tumours were isolated from a site outside the renal capsule.
CC	CC	RNA was isolated from primary tumours and metastases, reverse transcribed
CC	CC	and subjected to differential display PCR. The sequences were analysed to
CC	CC	obtain metastatic sequences, e.g. the sequence encoding the present
CC	CC	sequence. The method can be used to detect, diagnose and treat disorders
CC	CC	related to metastasis, or treat malignant or non-malignant disorders, can
CC	CC	e.g. hyperplasia, dysplasia and hypertrophy. The metastatic sequence, can

CC be used to screen a biological sample for metastasis, and it or its  
CC expression product may also be used to treat a metastatic disorder  
XX  
SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 5.7e-262;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDLVGNERSAYMRHNLQNGWSSDENDWNEKLYP 60  
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDLVGNERSAYMRHNLQNGWSSDENDWNEKLYP 60

Qy 61 VKKGDNRWNSKGGRRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120  
Db 61 VKKGDNRWNSKGGRRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180  
Db 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180

Qy 181 GOYFQKLGRCRSRVSVNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDVYVVTQIIPVFV 240  
Db 181 GOYFQKLGRCRSRVSVNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDVYVVTQIIPVFV 240

Qy 241 TMFQKNDNRNSDETFLKDLPTIMEDVLHDPSSHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300  
Db 241 TMFQKNDNRNSDETFLKDLPTIMEDVLHDPSSHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300

Qy 301 HTYVLNGTFSNLNTVKAAPGCPPPPPPPRPSKPTPSLGPAGNPLELSRIPDENQCIN 360  
Db 301 HTYVLNGTFSNLNTVKAAPGCPPPPPPPRPSKPTPSLGPAGNPLELSRIPDENQCIN 360

Qy 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPBESSLIDFVVTQCGSIPTVCITIIISDPT 420  
Db 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPBESSLIDFVVTQCGSIPTVCITIIISDPT 420

Qy 421 CEITQNTVCSFVDVDEMCLLTVRTFTNGSGTYCVNLTLGDDTSLALTSLISVDPDRPAS 480  
Db 421 CEITQNTVCSFVDVDEMCLLTVRTFTNGSGTYCVNLTLGDDTSLALTSLISVDPDRPAS 480

Qy 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540  
Db 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540

Qy 541 FPGNQEKDPLLNKQEFKGV 560  
Db 541 FPGNQEKDPLLNKQEFKGV 560

RESULT 2  
AAB11329  
ID AAB11329 standard; protein; 560 AA.  
XX  
AC AAB11329;  
XX  
AC AAB11329;  
XX  
DT 21-FEB-2001 (first entry)  
XX  
DE Human lung cancer-associated protein L528S.  
XX  
KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;  
XX vaccine; detection.  
XX  
OS Homo sapiens.  
XX  
FN WO200061612-A2. ✓  
XX  
PD 19-OCT-2000.  
XX  
FF 03-APR-2000; 2000WO-US008896.  
XX  
FF 02-APR-1999; 99US-00285479.  
PR 17-DEC-1999; 99US-00466396.  
PR

PR 30-DEC-1999; 99US-00476496.  
PR 10-JAN-2000; 2000US-00480884.  
PR 22-FEB-2000; 2000US-00510376.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Wang T, Fan L;  
XX  
XX WPI; 2000-628399/60.  
XX  
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor  
PT protein is used for detecting and monitoring progression of lung cancer  
PT in a patient.  
XX  
XX Claim 3; Page 203-204; 261pp; English.  
XX  
CC This invention describes a novel isolated polypeptide (I) which  
CC comprising an immunogenic portion of a lung tumor protein or variant (P2)  
CC which have cytostatic activity. The polypeptides and polynucleotides are  
CC used in compositions and vaccines to inhibit the development of cancer,  
CC especially lung cancer, in a patient. Methods described in the invention  
CC can be used to monitor the progression of a cancer by carrying out the  
CC detection at subsequent time points and comparing the results from the  
CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient  
CC are treated with P2, polynucleotides encoding P2 or antigen presenting  
CC cells expressing P2 and then administered to the patient to inhibit  
CC development of cancer  
XX  
SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 3; Length 560;  
Best Local Similarity 100.0%; Pred. No. 5.7e-262;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDLVGNERSAYMRHNLQNGWSSDENDWNEKLYP 60  
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDLVGNERSAYMRHNLQNGWSSDENDWNEKLYP 60

Qy 61 VKKGDNRWNSKGGRRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120  
Db 61 VKKGDNRWNSKGGRRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180  
Db 121 RNEAGLSADPYVYNWTAWSESDGNGTGQSHNVFPDGKPPHPGWRWNFIYVFTL 180

Qy 181 GOYFQKLGRCRSRVSVNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDVYVVTQIIPVFV 240  
Db 181 GOYFQKLGRCRSRVSVNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDVYVVTQIIPVFV 240

Qy 241 TMFQKNDNRNSDETFLKDLPTIMEDVLHDPSSHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300  
Db 241 TMFQKNDNRNSDETFLKDLPTIMEDVLHDPSSHFLNYSTINVKWSFGDNTGLFVSTNHTVN 300

Qy 301 HTYVLNGTFSNLNTVKAAPGCPPPPPPPRPSKPTPSLGPAGNPLELSRIPDENQCIN 360  
Db 301 HTYVLNGTFSNLNTVKAAPGCPPPPPPPRPSKPTPSLGPAGNPLELSRIPDENQCIN 360

Qy 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPBESSLIDFVVTQCGSIPTVCITIIISDPT 420  
Db 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPBESSLIDFVVTQCGSIPTVCITIIISDPT 420

Qy 421 CEITQNTVCSFVDVDEMCLLTVRTFTNGSGTYCVNLTLGDDTSLALTSLISVDPDRPAS 480  
Db 421 CEITQNTVCSFVDVDEMCLLTVRTFTNGSGTYCVNLTLGDDTSLALTSLISVDPDRPAS 480

Qy 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540  
Db 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540

Qy 541 FPGNQEKDPLLNKQEFKGV 560  
Db 541 FPGNQEKDPLLNKQEFKGV 560

```

RESULT 3
ABB78200
ID ABB78200 standard; protein; 560 AA.
XX
AC ABB78200;
XX
DT 25-NOV-2002 (first entry)
XX
DE Amino acid sequence of human HGFIN.
XX
KW Human; cell differentiation; white blood cell; bone marrow cell;
KW haematopoietic growth factor inducible neurokinin-1; HGFIN;
KW progenitor proliferation; acute myeloid leukemia; non-Hodgkin's disease;
KW acute lymphocytic leukemia; chronic myeloid leukemia;
KW chronic lymphocytic leukemia; Hodgkin's disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..485
FT /note= "extracellular domain"
FT Misc-difference 256
FT /note= "Leu encoded by CC"
FT Domain 485..508
FT /note= "transmembrane domain"
FT Domain 509..560
FT /note= "intracellular domain"
XX
PN WO200262947-A2.
XX
PD 15-AUG-2002.
XX
PF 20-OCT-2001; 2001WO-US050204.
XX
PR 20-OCT-2000; 2000US-0241881P.
XX
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
PI Pranela R;
XX
PI 2002-657531/70.
XX
DR N-PSDB; ABQ78551.
XX
XX
XX Hematopoietic growth factor inducible neurokinin-1 type polypeptide and
XX polynucleotide for treating a disease associated with abnormal bone
XX marrow cell differentiation or proliferation, e.g. leukemia.
XX
PS Claim 12; Page 123-125; 125pp; English.
XX
XX The present sequence represents haematopoietic growth factor inducible
XX neurokinin-1 type (HGFIN) polypeptide. HGFIN induces white blood cell
XX differentiation and inhibits progenitor proliferation. HGFIN polypeptides
XX and polynucleotides are useful for treating a disease associated with
XX abnormal bone marrow cell differentiation or proliferation, especially
XX acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid
XX leukemia, chronic lymphocytic leukemia, Hodgkin's and non-Hodgkin's
XX disease
XX
XX Sequence 560 AA;
XX
Query Match 100.0%; Score 3026; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MECLYFFLGFLLAARLPDAAKRFHDVLGNERPSAYMEHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYFFLGFLLAARLPDAAKRFHDVLGNERPSAYMEHNLQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSWKGGRVQAVLTSPALVGSNITFAVNLIPRCQKEDANGNIVYEKNC 120
DB 61 VVKRGDMRWKNSWKGGRVQAVLTSPALVGSNITFAVNLIPRCQKEDANGNIVYEKNC 120

```

```
XX PS Disclosure; Page 281-283; 374pp; English.
XX CC The present invention describes human lung tumour proteins. Human lung
XX CC tumour proteins and polynucleotides have cytostatic and immunostimulant
XX CC activities, and can be used in vaccine production. Compositions
XX CC comprising the lung tumour proteins, polynucleotides, antibodies, fusion
XX CC proteins, T cell populations, or antigen presenting cells that express
XX CC the lung tumour proteins are useful for treating lung cancer or
XX CC stimulating an immune response. ABL48959 to ABL49300 and ABL74946 to
XX CC ABL75070 represent sequences used in the exemplification of the present
XX CC invention
XX SQ Sequence 560 AA;
XX
XX Query Match 100.0%; Score 3026; DB 5; Length 560;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-262;
XX Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 180
QY 181 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 240
DB 181 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 240
QY 241 TMFQKNDNRNSDETFLKDLPIMFVDLIHDPGSHFLNYSNTYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRNSDETFLKDLPIMFVDLIHDPGSHFLNYSNTYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
QY 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
DB 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVDPDRPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVDPDRPAS 480
QY 481 PLRMANSLISVGCIAIFVTVISLIVYKHKYENPIENSPGNVVRKSLSVFLNRAKVP 540
DB 481 PLRMANSLISVGCIAIFVTVISLIVYKHKYENPIENSPGNVVRKSLSVFLNRAKVP 540
QY 541 PPGNQEKDPLLNQBFKGVS 560
DB 541 PPGNQEKDPLLNQBFKGVS 560
XX
XX RESULT 5
XX ID ABP61881
XX AC ABP61881 standard; protein; 560 AA.
XX XX
XX DT 07-OCT-2002 (first entry)
XX XX
XX DE Human lung cancer associated protein sequence SEQ ID NO:225.
XX KW Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.
XX OS Homo sapiens.
```

```
XX WO200247534-A2. ✓
XX PD 20-JUN-2002.
XX 30-NOV-2001; 2001WO-US047576.
XX 12-DEC-2000; 2000US-00735705.
XX 07-MAY-2001; 2001US-00850716.
XX 28-JUN-2001; 2001US-00897778.
XX (CORI-) CORIXA CORP.
XX Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
XX McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedwick TS;
XX Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;
XX WPI; 2002-583465/62.
XX Novel lung carcinoma polynucleotide sequences and polypeptides encoded by
XX the polynucleotides, useful in pharmaceutical compositions such as
XX vaccines and as markers to indicate the presence of lung cancer.
XX Example 1; Page 289-290; 381pp; English.
XX The present invention describes isolated human lung carcinoma
XX polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic
XX activity, and can be used in gene therapy and in vaccines. Compositions
XX comprising (I) or (II) can be used for stimulating an immune response in
XX a patient and for treating lung cancer in a patient. Oligonucleotides of
XX (I) can be used for detecting the presence of a cancer in a patient, by
XX obtaining a biological sample from the patient, contacting the biological
XX sample with the oligonucleotide, detecting in the sample an amount of
XX polynucleotide that hybridises to the oligonucleotide and comparing the
XX amount of polynucleotide that hybridises to the oligonucleotide to a
XX predetermined cut-off value, and determining the presence of a cancer in
XX the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.
XX vaccines. (I) is useful as a marker to indicate the presence or absence
XX of a cancer such as lung cancer. ABQ92145 to ABQ92496 and ABP61866 to
XX ABP61992 represent sequences used in the exemplification of the present
XX invention
XX Sequence 560 AA;
XX
XX Query Match 100.0%; Score 3026; DB 5; Length 560;
XX Best Local Similarity 100.0%; Pred. No. 5.7e-262;
XX Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIYVEKNC 120
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 180
DB 121 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 180
QY 181 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 240
DB 181 RNEAGLSADPVYVNTAWSESDSGENGTSQHNVFPDGKPPHPGWRNRNFIYVFTL 240
QY 241 TMFQKNDNRNSDETFLKDLPIMFVDLIHDPGSHFLNYSNTYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRNSDETFLKDLPIMFVDLIHDPGSHFLNYSNTYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCOIN 360
QY 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
DB 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCCTIISDPT 420
```

Db 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPSSLDIDFVVTCCQSIPTEVCTIISDPT 420  
 QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
 Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
 QY 481 PLRMANSALISVGCCLAFVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVF 540  
 Db 481 PLRMANSALISVGCCLAFVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVF 540  
 QY 541 PFGNQEKDPLLNQKQEFKGV 560  
 Db 541 PFGNQEKDPLLNQKQEFKGV 560

## RESULT 6

ABUS6592

ID ABUS6592 standard; protein; 560 AA.

XX AC

XX ABUS6592;

DT 02-APR-2003 (first entry)

DE Lung cancer-associated polypeptide #185.

XX Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

OS WO200286443-A2.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US012476.

XX 18-APR-2001; 2001US-0284770P.

XX 10-MAY-2001; 2001US-0290492P.

XX 09-NOV-2001; 2001US-0339245P.

XX 13-NOV-2001; 2001US-0350666P.

XX 29-NOV-2001; 2001US-0334370P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX PA

XX Aziz N, Murray R;

XX WPI; 2003-093161/08.

XX N-PSDB; ABX76321.

XX Detecting a lung cancer-associated transcript in a cell from a patient

XX for treating lung cancer, by contacting a biological sample from the

XX patient with a polynucleotide that exhibits increased or decreased

XX expression in lung cancer.

XX Claim 27; Page 328; 453pp; English.

XX The invention relates to a method for detecting a lung cancer-associated

XX transcript in a cell from a patient, comprising contacting a biological

XX sample from the patient with a polynucleotide that selectively hybridises

XX to a sequence that is at least 80 % identical to a gene that exhibits

XX increased or decreased expression in lung cancer samples. Lung cancer-

XX associated polynucleotides and polypeptides are used for identifying a

XX compound that modulates a lung cancer-associated polypeptide, for

XX inhibiting proliferation of a lung cancer-associated cell to treat lung

XX cancer in a patient and for treating a mammal having lung cancer by

XX administering a modulatory compound identified. The methods are useful

XX for treating lung cancer, such as small cell lung cancer, non-small cell

XX lung cancer or other benign or precancerous lesions, e.g. atelectasis,

CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
 CC for diagnostic purposes and as targets for screening for therapeutic  
 CC compounds that modulate lung cancer, such as antibodies. Sequences  
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
 CC invention

SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 6; Length 560;

Best Local Similarity 100.0%; Pred. No. 5.7e-262;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60

Db 1 MECLYYFLGFLLLAARLPDAAKRFHDLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60

QY 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIVYEKNC 120

Db 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIVYEKNC 120

QY 121 RNEAGLSADPYVYNWNTAWSESDGNGTQSHHNVFPDQKPPHHPGWRWRFVYVFHTL 180

Db 121 RNEAGLSADPYVYNWNTAWSESDGNGTQSHHNVFPDQKPPHHPGWRWRFVYVFHTL 180

QY 181 GOYFQKLGRCSVRVSVNTANVTGLPQMEVTVYRRHGRAYVPIAQKDVYVVTQDQIPVFV 240

Db 181 GOYFQKLGRCSVRVSVNTANVTGLPQMEVTVYRRHGRAYVPIAQKDVYVVTQDQIPVFV 240

QY 241 TMFQKQNRNSDETFLKOLPIMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300

Db 241 TMFQKQNRNSDETFLKOLPIMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300

QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360

Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360

QY 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPSSLDIDFVVTCCQSIPTEVCTIISDPT 420

Db 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPSSLDIDFVVTCCQSIPTEVCTIISDPT 420

QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDRPAS 480

Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTTCVNLTLGDDTSLALTSTLISVPRDRPAS 480

QY 481 PLRMANSALISVGCCLAFVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVF 540

Db 481 PLRMANSALISVGCCLAFVTVISLIVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVF 540

QY 541 PFGNQEKDPLLNQKQEFKGV 560

Db 541 PFGNQEKDPLLNQKQEFKGV 560

## RESULT 7

ABG72962

ID ABG72962 standard; protein; 560 AA.

XX AC

XX ABG72962;

DT 08-APR-2003 (first entry)

DE Human osteoactivin homologue, nmb.

XX Human; osteoactivin; osteopathic; antiinflammatory; vaccine;

XX bone cell differentiation stimulator; gene therapy; bone formation;

XX osteoporosis; periodontal disease; ectopic bone formation; osteopetrosis;

XX bone disorder; osteogenesis; enzyme.

XX Homo sapiens.

XX US2002151486-A1.

PN

XX 17-OCT-2002.  
XX 30-AUG-2001; 2001US-00943075.  
XX 30-AUG-2000; 2000US-0229006P.  
XX (POPO/) POPOFF S N.  
XX (SAFA/) SAFADI F F.  
XX (OWEN/) OWEN T A.  
XX (SMOC/) SMOCK S L.  
XX Popoff SN, Safadi FF, Owen TA, Smock SL;  
XX WPI; 2003-182528/18.  
XX Novel isolated osteoactivin protein and gene encoding the protein, useful  
XX for stimulating bone differentiation and for treating bone disorders  
XX including osteoporosis and periodontal disease.  
XX Claim 24; Fig 2B; 38pp; English.  
XX The invention describes an isolated and substantially pure osteoactivin  
XX protein (I) that stimulates bone cell differentiation. A therapeutic  
XX composition comprising (I), the polynucleotide (I) encoding (I) or a  
XX biologically active fragment of (I) is useful for stimulating bone  
XX formation in a mammal. The composition is also useful for treating a bone  
XX disorder including osteoporosis and periodontal disease. A second  
XX therapeutic composition comprising an anti-(I)-antibody or an agent that  
XX inhibits osteoactivin-mediated bone formation is useful for inhibiting  
XX formation in a mammal. The second composition is also useful for treating  
XX ectopic bone formation and osteopetrosis. (I) and (II) are also useful  
XX for developing novel therapeutic compositions for bone disorders, and for  
XX stimulating osteogenesis. The polynucleotide is also useful in gene  
XX therapy. This is the amino acid sequence of human nmb, a homologue of the  
XX rat osteoactivin of the invention used in the creation of antibodies for  
XX detection of nmb and osteoactivin  
XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 6; Length 560;  
Best Local Similarity 100.0%; Pred. No. 5.7e-262;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MECLYYFLGLLLAARLPDAAKPHDVLGNRPSPAYMREHNQNLGSSDNDWNEKLYP 60  
DB 1 MECLYYFLGLLLAARLPDAAKPHDVLGNRPSPAYMREHNQNLGSSDNDWNEKLYP 60  
QY 61 VWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIVYEKNC 120  
DB 61 VWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIVYEKNC 120  
QY 121 RNEAGLSADPYVYNTAWSESDSGENGTSQSHNVFPDGFPHHPGWRWNFTYVFTL 180  
DB 121 RNEAGLSADPYVYNTAWSESDSGENGTSQSHNVFPDGFPHHPGWRWNFTYVFTL 180  
QY 181 GQYFQKGRCSRVSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVVYVTTQIPVFV 240  
DB 181 GQYFQKGRCSRVSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVVYVTTQIPVFV 240  
QY 241 TMFQKDRNSDSEDTFLKDLPLIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TMFQKDRNSDSEDTFLKDLPLIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLNTVKAAPGCPPPPPPPRPSKPTSLGAGNPNLESLRIPDENCOIN 360  
DB 301 HTYVLNGTFSNLNTVKAAPGCPPPPPPPRPSKPTSLGAGNPNLESLRIPDENCOIN 360  
QY 361 RYGHFQATITVEGLEVNIQMTDVLMPWPWPESSLIIDFVTCQGSIPTEVCIIISDPT 420  
DB 361 RYGHFQATITVEGLEVNIQMTDVLMPWPWPESSLIIDFVTCQGSIPTEVCIIISDPT 420  
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLGDDTSLALTSTLISVDPDRPAS 480

DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLGDDTSLALTSTLISVDPDRPAS 480  
QY 481 PLRMANSALISVGCLAIFTVVISLLVYKKHKEYNPIENSPGVNRSGLSVFLNRKAVF 540  
DB 481 PLRMANSALISVGCLAIFTVVISLLVYKKHKEYNPIENSPGVNRSGLSVFLNRKAVF 540  
QY 541 FPGNQEKDPLLNQKQEPKGV 560  
DB 541 FPGNQEKDPLLNQKQEPKGV 560  
RESULT 8  
ID ABU70852  
XX ABU70852 standard; protein; 560 AA.  
XX AC ABU70852;  
XX DT 10-JUN-2003 (first entry)  
XX DE Human adipocyte Selected Interacting domain, SID, #483.  
XX KW Human; prey; adipocyte; SID; selected interacting domain; anorectic;  
XX KW anti-diabetic; protein-protein interaction; diabetes;  
XX KW yeast 2-hybrid assay; metabolic disorder; obesity.  
XX OS Homo sapiens.  
XX PN WO200286122-A2. ✓  
XX PD 31-OCT-2002.  
XX PF 14-MAR-2002; 2002WO-EP003768.  
XX PR 14-MAR-2001; 2001US-0275734P.  
XX PA (HYBR-) HYBRIGENICS.  
XX PI Legrain P, Daviet L;  
XX WPI; 2003-103412/09.  
XX N-PSDB; ACAS7396.  
XX New complex between two interacting proteins in adipocyte cells, useful  
XX for identifying selected interacting domains that modulate protein  
XX interactions, or for preventing or treating metabolic disorders such as  
XX obesity or diabetes.  
XX Claim 6; Page 269-270; 382pp; English.  
XX The invention relates to a complex between two interacting proteins in  
XX adipocyte cells, given in the specification. The proteins are identified  
XX by selecting a bait protein from a known adipocyte marker and then  
XX performing a yeast 2-hybrid selection to isolate prey proteins encoded by  
XX members of an adipocyte cDNA library. The proteins are designated SID  
XX (RTM) (selected interacting domains) proteins. Also included are a  
XX polynucleotide encoding a polypeptide in the adipocyte cells, a  
XX recombinant host cell expressing at least one of the interacting  
XX polypeptides of the complex, selecting a modulating compound in adipocyte  
XX cells, a SID (RTM) polypeptide comprising any of the 738 amino acid  
XX sequences given in the specification (including its fragment or variant),  
XX a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences  
XX given in the specification (including its fragment or variant), a vector  
XX comprising the SID (RTM) polynucleotide, a recombinant host cell  
XX comprising the vector, a protein chip comprising the polypeptides and a  
XX record comprising all or part of the data, listed in the specification.  
XX The complex, polypeptides, polynucleotides and compounds are useful for  
XX preventing or treating metabolic disorders such as obesity or diabetes.  
XX The polynucleotides are useful as probes or primers. The complex is  
XX particularly useful for identifying selected interacting domains (SID  
XX (RTM)) for screening drugs that modulate the protein interaction, thus  
XX exhibiting the therapeutic effect. The present sequence represents a SID  
XX (prey) protein of the invention

```

XX Sequence 560 AA;
SQ
Query Match 100.0%; Score 3026; DB 6; Length 560;
Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MECLYYFLGFLLLAARLPDAAKRPHDVLGNRPSPAYMREHNQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRPHDVLGNRPSPAYMREHNQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
DB 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNWTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 180
DB 121 RNEAGLSADPYVYNWTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 180
QY 181 GQYFQKLGRCSVRVSVNTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 240
DB 181 GQYFQKLGRCSVRVSVNTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 240
QY 241 TMFQKNDNRSSDETFELKDLPIIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRSSDETFELKDLPIIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLNTVKAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
DB 301 HTYVLNGTFSNLNTVKAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVYTCQGSIPTEVCTIISDPT 420
DB 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVYTCQGSIPTEVCTIISDPT 420
QY 421 CSITQNTVCSPVDVDEMCLLTVRRTPFNGSGTYCVNLTLGDDTSLATLSLISVPRDRPAS 480
DB 421 CSITQNTVCSPVDVDEMCLLTVRRTPFNGSGTYCVNLTLGDDTSLATLSLISVPRDRPAS 480
QY 481 PLRMANSALISVGLAIFVTVISLVLYKHKKEYNPINSPGNVRSKGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGLAIFVTVISLVLYKHKKEYNPINSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQBFKGV 560
DB 541 PFGNQEKDPLLKNQBFKGV 560
RESULT 9
ADA28315
ID ADA28315 standard; protein; 560 AA.
XX
AC ADA28315;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human lung tumour L528S protein.
XX
KW cancer; lung cancer; gene therapy; vaccine; human;
KW lung squamous cell carcinoma.
XX
OS Homo sapiens.
XX
PN US2003064947-A1.
XX
PD 03-APR-2003.
XX
PF 30-NOV-2001; 2001US-00007700.
XX
PR 18-MAR-1998; 98US-00040802.
PR 27-JUL-1998; 98US-00123912.
PR 22-DEC-1998; 98US-00221107.
PR 02-APR-1999; 99US-00285479.
```

```

PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PR 04-APR-2000; 2000US-00542615.
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX (CORI-) CORIXA CORP.
PA
Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
Carter D, Watanabe Y, Peckham DW, Cai F, Poy TW,
WPI; 2003-540798/51.
XX
New isolated polynucleotides and polypeptides useful for diagnosing,
preventing and/or treating cancer, particularly lung cancer.
Example 7; Page 194-195; 296pp; English.
XX
The invention describes isolated polynucleotides and polypeptides useful
for diagnosing, preventing and/or treating cancer, particularly lung
cancer. A new isolated polynucleotide comprises: any of the 22 fully
defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
specification; complements of the nucleotide sequences cited above; at
least 10 contiguous residues of the nucleotide sequences cited above; a
sequence that hybridise to any of the nucleotide sequences under highly
stringent conditions; a sequence that is at least 75 or 90% identical to
the above nucleotide sequences; or degenerate variants of the above
nucleotide sequences. The composition and methods are useful in
diagnosing, preventing and/or treating cancer, particularly lung cancer,
in gene therapy and in vaccines. This is the amino acid sequence encoded
by a human lung tumour cDNA isolated from a lung squamous cell carcinoma
that may be useful in the diagnosis and treatment of lung cancer and
other disorders.
XX Sequence 560 AA;
QY Query Match 100.0%; Score 3026; DB 7; Length 560;
DB Best Local Similarity 100.0%; Pred. No. 5.7e-262;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MECLYYFLGFLLLAARLPDAAKRPHDVLGNRPSPAYMREHNQNGWSSDENDWNEKLYP 60
DB 1 MECLYYFLGFLLLAARLPDAAKRPHDVLGNRPSPAYMREHNQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
DB 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPRCQKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNWTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 180
DB 121 RNEAGLSADPYVYNWTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 180
QY 181 GQYFQKLGRCSVRVSVNTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 240
DB 181 GQYFQKLGRCSVRVSVNTANVTGTPQMEVTVYRRHGRAYVPIAQKDVYVVTQDIPVVF 240
QY 241 TMFQKNDNRSSDETFELKDLPIIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRSSDETFELKDLPIIMFDVLIHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLNTVKAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
DB 301 HTYVLNGTFSNLNTVKAAPGCPPPPPRPSKPTPSLGPAGDNPFLSLRIPDENCQIN 360
```



Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPRESSLIIDFVVTQGGSIPTVCIIISDPT 420  
 Db |||||  
 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPRESSLIIDFVVTQGGSIPTVCIIISDPT 420  
 Qy 421 CEITONTVCSPVDVDEMCLLTVRRFTFNGSGTYCVNLTLGGDTSLALSTLISVPDRDPAS 480  
 Db |||||  
 421 CEITONTVCSPVDVDEMCLLTVRRFTFNGSGTYCVNLTLGGDTSLALSTLISVPDRDPAS 480  
 Qy 481 PLRMANSALISVGCIAIFVTVISLLVYKKHKEYNPINSPGNVRSKGLSVFLNRAKAVF 540  
 Db |||||  
 481 PLRMANSALISVGCIAIFVTVISLLVYKKHKEYNPINSPGNVRSKGLSVFLNRAKAVF 540  
 Qy 541 FPGNQEKDPLLNQEFKGV 560  
 Db |||||  
 541 FPGNQEKDPLLNQEFKGV 560

RESULT 10  
 ADE95620  
 ID ADE95620 standard; protein; 560 AA.  
 XX  
 AC ADE95620;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Human NOVX28c protein.  
 XX  
 KW NOVX protein; biochemical stimulation; physiological stimulation;  
 KW cardiant; antiarteriosclerotic; hypotensive; cytostatic; anorectic;  
 KW antirheumatic; antiarthritic; antidiabetic; nephrotropic; dermatological;  
 KW immunosuppressive; anti-HIV; antiinflammatory; neuroprotective;  
 KW neutropic; antipeptidic; antiparkinsonian; antiasthmatic; neuroleptic;  
 KW antidepressant; anti-allergic; gynaecological; gene therapy; vaccine;  
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; hypertension;  
 KW cancer; obesity; rheumatoid arthritis; diabetes; glomerulonephritis;  
 KW psoriasis; skin disorder; AIDS; inflammation; multiple sclerosis;  
 KW Alzheimer's disease; Parkinson's disease; asthma; schizophrenia;  
 KW depression; allergy; fertility disorder; NOVX28c.  
 XX  
 OS Homo sapiens.  
 XX

Key Location/Qualifiers  
 FH Misc-difference 47  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Arg as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 57  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Glu as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 111  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Asp as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 130  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Pro as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 154  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Ser as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 162  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Leu as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 294  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Phe as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 324  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Leu as a result of a

FT single nucleotide polymorphism"  
 FT Misc-difference 460  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Asp as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 487  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Ser as a result of a  
 FT single nucleotide polymorphism"  
 FT Misc-difference 526  
 FT /label= OTHER  
 FT /note= "OTHER= May be substituted by Arg as a result of a  
 FT single nucleotide polymorphism"  
 FT  
 PN WO2003050245-A2.  
 XX  
 XX 19-JUN-2003.  
 XX  
 XX 03-DEC-2002; 2002WO-US038594.  
 XX  
 XX 05-DEC-2001; 2001US-0336600P.  
 XX 07-DEC-2001; 2001US-0338285P.  
 XX 12-DEC-2001; 2001US-0341346P.  
 XX 17-DEC-2001; 2001US-0341477P.  
 XX 17-DEC-2001; 2001US-0341540P.  
 XX 20-DEC-2001; 2001US-0342592P.  
 XX 27-DEC-2001; 2001US-0344297P.  
 XX 31-DEC-2001; 2001US-0344903P.  
 XX 17-APR-2002; 2002US-0373288P.  
 XX 15-MAY-2002; 2002US-0380981P.  
 XX 17-MAY-2002; 2002US-0381495P.  
 XX 28-MAY-2002; 2002US-0383534P.  
 XX 28-MAY-2002; 2002US-0383744P.  
 XX 29-MAY-2002; 2002US-0383829P.  
 XX 29-MAY-2002; 2002US-0384024P.  
 XX 07-AUG-2002; 2002US-0401788P.  
 XX 26-AUG-2002; 2002US-0406353P.  
 XX 31-OCT-2002; 2002US-00401788.  
 XX 02-DEC-2002; 2002US-00406353.  
 XX  
 FA (CURA-) CURAGEN CORP.  
 XX  
 XX Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Chillakuru RA;  
 PI Edinger SR, Gerlach VL, Gorman L, Gould-Rothberg BE, Guo X;  
 PI Jeffers WE, Ji W, Li L, Malyankar UM, Miller CE, Murphey R;  
 PI Patturajan M, Peyman JA, Rastelli L, Rieger DK, Shenoy SG;  
 PI Smithson G, Starling G, Taupier RJ, Voss EZ, Zhong H, Zhong M;  
 XX  
 XX WPI; 2003-513974/48.  
 DR N-PSDB; ADE95619.  
 XX  
 XX New NOVX polypeptides and nucleic acids, useful for preventing or  
 FT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,  
 FT atherosclerosis or diabetes, and in chromosome mapping, tissue typing or  
 FT pharmacogenomics.  
 XX  
 XX Claim 2; SEQ ID NO 152; 211pp; English.  
 XX  
 XX This invention relates to novel NOVX proteins, and the DNA sequence which  
 CC encode them, having properties related to stimulation of biochemical or  
 CC physiological responses in a cell, a tissue, an organ or an organism.  
 CC Compounds which modulate the proteins of the invention may have cardiac,  
 CC antiarteriosclerotic, hypotensive, cytostatic, anorectic, antirheumatic,  
 CC antiarthritic, antidiabetic, nephrotropic, dermatological,  
 CC immunosuppressive, anti-HIV, antiinflammatory, neuroprotective,  
 CC neutropic, antipeptidic, antiparkinsonian, antiasthmatic, neuroleptic,  
 CC antidepressant, anti-allergic or gynaecological activities. The DNA  
 CC sequences of the invention may be useful for gene therapy whilst the  
 CC protein sequences may allow the development of a vaccine. The protein is  
 CC useful in the manufacture of a medicament for treating a syndrome  
 CC associated with a human disease. The invention may be useful in  
 CC diagnosing, treating or preventing NOVX-associated disorders, for example  
 CC cardiomyopathy, atherosclerosis, hypertension, cancer, obesity,



CC rheumatoid arthritis, diabetes, glomerulonephritis, psoriasis, skin  
 CC disorders, AIDS, inflammation, multiple sclerosis, Alzheimer's disease,  
 CC Parkinson's disease, asthma, schizophrenia, depression, allergies or  
 CC fertility disorders. The nucleic acids may further be used as  
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive  
 CC medicine, and pharmacogenomics. The present sequence is the amino acid  
 CC sequence of the human NOVX28c protein of the invention.

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 5,7e-262;  
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
 Db 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
 QY 61 VVKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIYVEKNC 120  
 Db 61 VVKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIYVEKNC 120  
 QY 121 RNEAGLSADPYNNWTANSESDGNGTQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180  
 Db 121 RNEAGLSADPYNNWTANSESDGNGTQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180  
 QY 181 GOYFQKLGRCVRSVNTANVTLGQPMVTVYRRHGRAYVPIAQVKDYVVTDOIIPV 240  
 Db 181 GOYFQKLGRCVRSVNTANVTLGQPMVTVYRRHGRAYVPIAQVKDYVVTDOIIPV 240  
 QY 241 TFMQKNDNRNSDETFLKDLPIFMDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
 Db 241 TFMQKNDNRNSDETFLKDLPIFMDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
 QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLESRIIPDENQCIN 360  
 Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLESRIIPDENQCIN 360  
 QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420  
 Db 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420  
 QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVPDRDPAS 480  
 Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVPDRDPAS 480  
 QY 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540  
 Db 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540  
 QY 541 PFGNQEKDPLLKNQEFKGV 560  
 Db 541 PFGNQEKDPLLKNQEFKGV 560

RESULT 11

ADH36879

ID ADH36879 standard; protein; 560 AA.

XX

AC ADH36879;

XX 11-MAR-2004 (first entry)

XX Human lung cancer-related protein #16.

KW lung cancer; lung tumour; immune response stimulation;  
 KW tumour protein specific T cell; human.

XX Homo sapiens.

XX WO2003086175-A2.

XX 23-OCT-2003.

XX  
 PF  
 XX  
 PR  
 PR  
 XX  
 PA  
 XX  
 PI  
 PI  
 PI  
 XX  
 DR  
 XX  
 PT  
 PT  
 PT  
 PT  
 XX  
 PS  
 XX  
 CC  
 CC  
 CC  
 CC  
 CC  
 CC  
 CC  
 CC  
 SQ

07-APR-2003; 2003WO-US010945.

05-APR-2002; 2002US-00117982.

04-DEC-2002; 2002US-00313986.

(CORI-) CORIXA CORP.

Mericle B, Fanger GR, Vedvick TS, Carter D, Watanabe Y;

Henderson RA, Kalos MD, Spies GA, Foy TM, Wang T, McNabb A;

Reed SG;

WPI; 2003-845247/78.

Use of compositions comprising lung tumor polypeptides or

polynucleotides, for inhibiting lung cancer or tumor progression,

stimulating immune response, or stimulating and/or expanding T cells

specific for a tumor protein.

Disclosure; SEQ ID NO 225; 456pp; English.

The invention comprises a method for inhibiting the development of lung

cancer, the method involves the use of compositions which contain lung

tumor polypeptides/polynucleotides. The methods and compositions of the

invention are useful for inhibiting lung cancer development or tumor

progression, stimulating immune response, or stimulating and/or expanding

T cells specific for a tumor protein. The present amino acid sequence

represents a human lung cancer-related protein.

Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;

Best Local Similarity 100.0%; Pred. No. 5,7e-262;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60

Db 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60

QY 61 VVKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIYVEKNC 120

Db 61 VVKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIIPRCOKEDANGNIYVEKNC 120

QY 121 RNEAGLSADPYNNWTANSESDGNGTQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180

Db 121 RNEAGLSADPYNNWTANSESDGNGTQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180

QY 181 GOYFQKLGRCVRSVNTANVTLGQPMVTVYRRHGRAYVPIAQVKDYVVTDOIIPV 240

Db 181 GOYFQKLGRCVRSVNTANVTLGQPMVTVYRRHGRAYVPIAQVKDYVVTDOIIPV 240

QY 241 TFMQKNDNRNSDETFLKDLPIFMDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300

Db 241 TFMQKNDNRNSDETFLKDLPIFMDVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300

QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLESRIIPDENQCIN 360

Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGDNPLESRIIPDENQCIN 360

QY 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420

Db 361 RYGHFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVCTIISDPT 420

QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVPDRDPAS 480

Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVPDRDPAS 480

QY 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540

Db 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPPIENSPGNVRSKGLSVFLNRAKAV 540

QY 541 PFGNQEKDPLLKNQEFKGV 560

Db 541 PFGNQEKDPLLKNQEFKGV 560



XX Disclosure; SEQ ID NO 7; 342pp; English.

XX The invention relates to a compound comprising a binding moiety which selectively binds to a protein or polypeptide listed in the specification (e.g. human autotaxin polypeptide or human CD24 signal transducer polypeptide), and a further moiety. The compound is useful in medicine or in the treatment, imaging, diagnosis or prognosis of mantle cell lymphomas (MCL). It is used in preparing a medicament for treating MCL, a diagnostic or prognostic agent for MCL, or an agent for imaging MCL cells in the body of an individual. This sequence corresponds to one of the polypeptides of the invention.

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;  
Best Local Similarity 100.0%; Pred. No. 5.7e-262;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60  
DB 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60

QY 61 VMKRGDMRWKNSKGRVQAVLTS DSPALVGSNITFAVNLIPRCQKEDANGNIYVKNC 120  
DB 61 VMKRGDMRWKNSKGRVQAVLTS DSPALVGSNITFAVNLIPRCQKEDANGNIYVKNC 120

QY 121 RNEAGLSADPYVYNTANVTLSGQSHHNPDPCKPPHPGWRNRNFIYVFHTL 180  
DB 121 RNEAGLSADPYVYNTANVTLSGQSHHNPDPCKPPHPGWRNRNFIYVFHTL 180

QY 181 GQYFQKLGRCVRSVNTANVTLSGQSHHNPDPCKPPHPGWRNRNFIYVFHTL 240  
DB 181 GQYFQKLGRCVRSVNTANVTLSGQSHHNPDPCKPPHPGWRNRNFIYVFHTL 240

QY 241 TFWQKNDNRSSDETFLKOLPIMFDVLIHDPSPHFLNYSNTYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TFWQKNDNRSSDETFLKOLPIMFDVLIHDPSPHFLNYSNTYKWSFGDNTGLFVSTNHTVN 300

QY 301 HTYVLNGTFLSLNTVKAAPGCPPPPPPPPPSKPTPSLGPAGDNPLELSRIPDENCQIN 360  
DB 301 HTYVLNGTFLSLNTVKAAPGCPPPPPPPPPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPPESSLIDFVVTCCQSIPTEVCTIISDPT 420  
DB 361 RYGHFOATITIVGILEVNIQMTDVLMPVPPESSLIDFVVTCCQSIPTEVCTIISDPT 420

QY 421 CEITQNTVCSYVDVDEMCLLTVRTFNGSGTTCVNLITGDDTSLALTSLISVPPDRPAS 480  
DB 421 CEITQNTVCSYVDVDEMCLLTVRTFNGSGTTCVNLITGDDTSLALTSLISVPPDRPAS 480

QY 481 PLRMANSALISVGCILAIFVTVISLVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGCILAIFVTVISLVYKHKYNIPIENSPGNVRSKGLSVFLNRAKAVF 540

QY 541 PFGNQEKDPLLKNQBFKGS 560  
DB 541 PFGNQEKDPLLKNQBFKGS 560

RESULT 14  
ADN39940  
ID ADN39940 standard; protein; 560 AA.

XX ADN39940;

XX 17-JUN-2004 (first entry)

XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C310.

XX Human; differential expression; cancer; angiogenic disorder;  
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
KW inflammatory disease; autoimmune disease;

KW retinal neovascularisation syndrome; scarring; uterine fibroid;  
KW detection; diagnosis; prognosis; drug screening; drug targeting;  
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;  
KW vulnery; gene therapy; vaccine.  
OS Homo sapiens.  
PN WO2003042661-A2.  
PD 22-MAY-2003.  
XX 13-NOV-2002; 2002WO-US036810.  
XX 13-NOV-2001; 2001US-0350666P.  
PR 21-NOV-2001; 2001US-0332464P.  
PR 29-NOV-2001; 2001US-0334393P.  
PR 03-DEC-2001; 2001US-0335394P.  
PR 14-DEC-2001; 2001US-0340376P.  
PR 08-JAN-2002; 2002US-0347211P.  
PR 10-JAN-2002; 2002US-0347349P.  
PR 08-FEB-2002; 2002US-0355250P.  
PR 13-FEB-2002; 2002US-0356714P.  
PR 20-FEB-2002; 2002US-0359077P.  
PR 29-MAR-2002; 2002US-036809P.  
PR 04-APR-2002; 2002US-0370110P.  
PR 12-APR-2002; 2002US-0372246P.  
PR 05-JUN-2002; 2002US-0386614P.  
PR 16-JUL-2002; 2002US-0396839P.  
PR 22-JUL-2002; 2002US-039775P.  
PR 22-JUL-2002; 2002US-0397845P.  
PR 09-SEP-2002; 2002US-0409450P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevesi PA;  
PI Mack DH, Murray R, Watson SR, Wilson KE, Zioltnik A;  
XX WPI; 2003-468649/44.  
DR N-PSDB; ADN39723.

XX Determining the presence or absence of a pathological cell in a patient.  
XX useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.

Claim 12; SEQ ID NO C310; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a polypeptide of the invention.

XX Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 7; Length 560;  
Best Local Similarity 100.0%; Pred. No. 5.7e-262;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60  
DB 1 MECLYFGLFLLAARLPDAAKRFHDVILGNRPSPAYMREHNLQNGWSSDENDWNEKLYP 60

Qy 61 VKKGDMDWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120  
Db |||||  
Qy 61 VKKGDMDWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120  
Db |||||  
Qy 121 RNEAGLSADPYVYNTWTAWSESDSGENGTCGSHHNVFPDGKPPHPGRRWNFIYVFTL 180  
Db |||||  
Qy 121 RNEAGLSADPYVYNTWTAWSESDSGENGTCGSHHNVFPDGKPPHPGRRWNFIYVFTL 180  
Db |||||  
Qy 181 GQYFQKLGRCVRSVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVVTDOIPVFV 240  
Db |||||  
Qy 181 GQYFQKLGRCVRSVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVVTDOIPVFV 240  
Db |||||  
Qy 241 TMFQKNDNRNSDETFLKDLPIPMFDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300  
Db |||||  
Qy 241 TMFQKNDNRNSDETFLKDLPIPMFDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300  
Db |||||  
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360  
Db |||||  
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360  
Db |||||  
Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420  
Db |||||  
Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420  
Db |||||  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSALTSTLISVDPDRPAS 480  
Db |||||  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSALTSTLISVDPDRPAS 480  
Db |||||  
Qy 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db |||||  
Qy 481 PLRMANSALISVGCLAIPTVTVISLLVYKKHKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db |||||  
Qy 541 PFGNQEKDPLLNQEFKGV 560  
Db |||||  
Qy 541 PFGNQEKDPLLNQEFKGV 560  
Db |||||

## RESULT 15

ADH56342  
ID ADH56342 standard; protein; 560 AA.

XX  
AC ADH56342;

DT 25-MAR-2004 (first entry)

XX  
DE Human nmb protein SEQ ID NO:1.

XX  
KW respiratory disease; human; nmb; antiinflammatory; antiaethmatic;  
KW antiallergic; neuroprotective; immunosuppressive; antidiabetic;  
KW antiarthritic; dermatological; gastrointestinal; antirheumatic; asthma;  
KW hayfever; chronic bronchitis; chronic obstructive lung disease;  
KW immune disorder; multiple sclerosis; Sjogren's disease;  
KW insulin-resistant diabetes; rheumatoid arthritis; lupus erythematosus;  
KW atopic dermatitis; irritable bowel disease.

XX  
OS Homo sapiens.

XX  
FN WO2004002516-A1. ✓

XX  
PD 08-JAN-2004.

XX  
PP 27-JUN-2003; 2003WO-JP008168.

XX  
PR 28-JUN-2002; 2002JP-00190790.

XX  
PA (TAKE ) TAKEDA CHEM IND LTD.

XX  
PI Nakanishi A, Iwashita H, Morita S, Matsumoto T, Yamasaki M;

XX  
DR WPI; 2004-082883/08.

XX  
DR N-PSDB; ADH56343.

PT Drug compositions containing inhibitor of human nmb protein activity for  
PT treatment and prevention of respiratory diseases including asthma and  
XX chronic obstructive lung disease.

PS Claim 1; SEQ ID NO 1; 104pp; Japanese.

XX The present invention describes drug compositions (1) for prevention and  
CC treatment of respiratory diseases, which contain a compound or its salts  
CC inhibiting the activity of human nmb protein or its partial peptides or a  
CC protein of substantially similar sequence and activity, or inhibiting the  
CC expression of the gene encoding this protein. Also described: (1)  
CC antisense polynucleotides to all or part of the nmb gene, and drug  
CC compositions containing them; (2) antibodies to all or part of nmb  
CC protein, and drug compositions and diagnostic reagents containing them;  
CC (3) diagnostic reagents containing polynucleotides encoding all or part  
CC of nmb protein; (4) screening method for compounds for use in prevention  
CC and treatment of respiratory diseases, using nmb protein or its partial  
CC peptides, or using polynucleotides encoding all or part of nmb protein,  
CC or using a tobacco smoking induced chronic obstructive lung disease model  
CC mouse or an elastase-induced chronic obstructive lung disease model mouse  
CC ; (5) kits for the screening method; (6) compounds identified by the  
CC screening method; (7) a method for prevention and treatment of  
CC respiratory diseases using the identified compounds; and (8) drug  
CC compositions for prevention and treatment of respiratory diseases  
CC containing a compound or its salts inhibiting the binding activity of  
CC heparan sulfate proteoglycan. (1) have antiinflammatory, antiaethmatic,  
CC antiallergic, neuroprotective, immunosuppressive, antidiabetic,  
CC antiarthritic, dermatological, gastrointestinal, and antirheumatic  
CC activities. (1) can be used for the prevention, treatment and diagnosis  
CC of respiratory diseases such as acute asthma, chronic asthma, hayfever,  
CC chronic bronchitis and chronic obstructive lung disease, and other  
CC diseases including immune disorders such as multiple sclerosis, Sjogren's  
CC disease, insulin-resistant diabetes, rheumatoid arthritis, lupus  
CC erythematosus, atopic dermatitis, and irritable bowel disease. The  
CC present sequence represents human nmb, which is used in the  
CC exemplification of the present invention.

XX  
SQ Sequence 560 AA;

Query Match 100.0%; Score 3026; DB 8; Length 560;

Best Local Similarity 100.0%; Pred. No. 5.7e-262;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGLLLAARLPDAAKRFHDVIGNERPSAYMREHNQNLNGSSDENDNEKLYP 60

Db |||||

Qy 61 VWKRGDMRWKNKWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120

Db |||||

Qy 121 RNEAGLSADPYVYNTWTAWSESDSGENGTCGSHHNVFPDGKPPHPGRRWNFIYVFTL 180

Db |||||

Qy 181 GQYFQKLGRCVRSVSVNTANVTLPQMEVTVYRRHGRAYVPIAQVKDVIYVVTDOIPVFV 240

Db |||||

Qy 241 TMFQKNDNRNSDETFLKDLPIPMFDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300

Db |||||

Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

Db |||||

Qy 361 RYGHFOATITIVGILEVNIQMTDVLMPVWPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420

Db |||||

Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSALTSTLISVDPDRPAS 480

Db |||||

Db 421 CEITQNTVCSPVDVDEMCLITVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
 Qy 481 PLRMANSALISVGCLAIFVTVISLLVYKHKHKNPIENSPGNVVRSGLSVFLNRAKAVF 540  
 Db 481 PLRMANSALISVGCLAIFVTVISLLVYKHKHKNPIENSPGNVVRSGLSVFLNRAKAVF 540  
 Qy 541 FPGNQEKDPLLKNQEPKGV 560  
 Db 541 FPGNQEKDPLLKNQEPKGV 560

Search completed: December 24, 2005, 04:31:49  
 Job time : 197 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:25:27 ; Search time 42 Seconds  
(without alignments)  
1282.890 Million cell updates/sec

Title: US-09-853-880A-17  
Perfect score: 3026  
Sequence: 1 MECLYFGLFLLAARLPD.....PFGNQKDPLLKNQEPKGV 560

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 80: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	2	I38065
2	659	21.8	626	2	S53871
3	622.5	20.6	662	2	I38400
4	617.5	20.4	668	2	A41234
5	459	15.2	491	2	A43179
6	136	4.5	926	1	A41105
7	120.5	4.0	446	2	T07907
8	114.5	3.8	1009	2	C64483
9	111.5	3.7	457	2	I55976
10	111.5	3.7	555	2	S21766
11	109.5	3.6	348	2	AB3260
12	108	3.6	3848	2	T23694
13	106.5	3.5	848	2	T17414
14	106	3.5	1874	1	JQ0533
15	105	3.5	492	2	C96521
16	104.5	3.5	363	2	G83306
17	104.5	3.5	624	1	I51581
18	104.5	3.5	837	2	C69200
19	104.5	3.5	840	2	B69216
20	104.5	3.5	4302	2	A38971
21	104	3.4	913	2	T18503
22	102.5	3.4	674	2	A10922
23	102.5	3.4	921	2	A33718
24	102.5	3.4	979	2	A35913
25	101.5	3.4	445	2	D81716
26	101	3.3	588	2	T45564
27	101	3.3	705	2	T16088
28	101	3.3	2014	2	T21560
29	100.5	3.3	221	2	T07176

30	100	3.3	432	2	AB2222	twitching motility
31	100	3.3	13055	2	T16580	hypothetical prote
32	99.5	3.3	393	2	T33103	lin-1 protein - Ca
33	99.5	3.3	826	2	G90283	hypothetical prote
34	99.5	3.3	1165	2	I40644	botulinum neurotox
35	99.5	3.3	1621	2	T15264	hypothetical prote
36	99.5	3.3	26926	1	I38344	titin, cardiac mus
37	99	3.3	281	2	T23150	hypothetical prote
38	98.5	3.3	594	2	I58386	receptor tyrosine
39	98.5	3.3	594	2	I56248	receptor tyrosine
40	98.5	3.3	712	2	T18284	hypothetical prote
41	98.5	3.3	1048	2	T30815	platelet-derived g
42	98	3.2	673	1	HJSCDR	ATP-dependent DNA
43	98	3.2	4199	2	S76412	hypothetical prote
44	97.5	3.2	845	2	T40955	hypothetical prote
45	97.5	3.2	1193	2	E69933	hypothetical prote

ALIGNMENTS

RESULT 1

I38065  
gene NMB protein - human  
C:Species: Homo sapiens (man)  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
C:Accession: I38065  
R:Wetman, M.A.; Ajubi, N.; van Dinter, I.M.; Degen, W.G.; van Muijen, G.N.; Ruitter, I.  
Int. J. Cancer 60, 73-81, 1995  
A:Title: nmb, a novel gene, is expressed in low-metastatic human melanoma cell lines an  
A:Reference number: I38065; MUID:95113576; PMID:7814155  
A:Accession: I38065  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-560 <RES>  
A:Cross-references: UNIPROT:Q14956; UNIPARC:UPI0000047E2A; EMBL:X76534; NID:g666042; PI  
C:Genetics:  
A:Gene: GDB:NMB  
A:Cross-references: GDB:I20237; OMIM:162340  
A:Map position: 15q22-15qter

Query Match		100.0%	Score 3026;	DB 2;	Length 560;
Best Local Similarity		100.0%	Pred. No. 5.4e-228;	Mismatches 0;	Indels 0; Gaps 0;
Matches 560;		Conservative 0;			
QY	1	MECLYFGLFLLAARLPDAAKRFHDVLGNERPSAYMREHNLQNGWSSDENDWNEKLYP	60		
DB	1	MECLYFGLFLLAARLPDAAKRFHDVLGNERPSAYMREHNLQNGWSSDENDWNEKLYP	60		
QY	61	VWKGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIVYEKNC	120		
DB	61	VWKGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIVYEKNC	120		
QY	121	RNEAGLSADPYVYNNATSESDSGNGTCQSHHNPDPCKPFPHPGRWRFVYFHTL	180		
DB	121	RNEAGLSADPYVYNNATSESDSGNGTCQSHHNPDPCKPFPHPGRWRFVYFHTL	180		
QY	181	GOYFQKLGRCVSRVSVNTANVTGLPQLMETVYRRHGRAYVPIAOKVQVYVVTQDQIPVFV	240		
DB	181	GOYFQKLGRCVSRVSVNTANVTGLPQLMETVYRRHGRAYVPIAOKVQVYVVTQDQIPVFV	240		
QY	241	TMFQKNDNRNSDETFLKOLPIMFDVLIHDPSPHFLANYSTINYKWSFGDNTGLFVSTNHTVN	300		
DB	241	TMFQKNDNRNSDETFLKOLPIMFDVLIHDPSPHFLANYSTINYKWSFGDNTGLFVSTNHTVN	300		
QY	301	HTYVLNGTFSNLTKAAAPGCPPPPSPKTPSLGPDAGNPLELSRIPDENCOIN	360		
DB	301	HTYVLNGTFSNLTKAAAPGCPPPPSPKTPSLGPDAGNPLELSRIPDENCOIN	360		
QY	361	RYGHFQATITVEGILEVNIIOMTDVLMPVPPESSLIIDFVVTCCGSIPTVCTIISDPT	420		
DB	361	RYGHFQATITVEGILEVNIIOMTDVLMPVPPESSLIIDFVVTCCGSIPTVCTIISDPT	420		



```
Qy 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVPRDPAS 480
Db 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVPRDPAS 480
Qy 481 PLRMANSALISVGCLAIPTVTISLLVYKHKHEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
Db 481 PLRMANSALISVGCLAIPTVTISLLVYKHKHEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
Qy 541 PPGNQEKDPLLNQBFKGV 560
Db 541 PPGNQEKDPLLNQBFKGV 560

RESULT 2
Emel 17 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S53871
R:Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamaneni, C.; Bennett, D.; Pick
Nucleic Acids Res. 23, 154-158, 1995
A:Title: Mouse silver mutation is caused by a single base insertion in the putative cyto
A:Reference number: S53871; MUID:95175358; PMID:7870580
A:Accession: S53871
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-626 <KWO>
A:Cross-references: UNIPROT:Q60696; UNIPARC:UPI000002822F; GB:U14133; NID:g887940; PIDN:

Query Match 21.8%; Score 659; DB 2; Length 626;
Best Local Similarity 27.2%; Pred. No. 2.4e-43;
Matches 183; Conservative 99; Mismatches 211; Indels 180; Gaps 20;

Qy 7 FLGFLLLAARLPDAA--AKRFDVLGNRPSPAYMREHNQNLGWSDDNDWNEKLYPVWKR 64
Db 9 FLPLVLVALLAVGALEGRNQDLGVPRQLV-----TKTWNRQLYPEWT- 53

Qy 65 GDMWKNSWKGRVQAVLTSDSPALVGSNTITFAVNLIPPRCKEDANGNIYVEKN-CRNE 123
Db 54 -EVQGSNCWRGGQVSLRVINDPTLVGNASFSIALHPFGSQKVLPGQVWANNITNG 112

Qy 124 AGLSADPVVYNTWTAMSEDSDCGENTGQSHHNVFPDGKPPFPHHPGRRWNFIYVFHTLQY 183
Db 113 SQVWGQGVYP-----QEPDDA-----CVFPDGGPCSPGPKPKRSFVYVWKTGKY 159

Qy 184 FQKLRCSVRVSVNTANVTLPQLMEVTVYRRHG--RAYVPIAQKDVVYVTDQIPVFTM 242
Db 160 WQVLGGPVRSRSTIATRHAKLGTHTMEVTVYHRRGSQSVPLAHASSTFTITDQVPFVS 219

Qy 243 FQKNDRSSDETFLKDLPIMPDVLIHDPSPHFLANTSTINYKWSFGDNTGLFVSTNHTV 302
Db 220 SQLQALDGETKHFLENHPLIFALQLHDPGVLAEADLSYTWDFGDTGLISRALDVHT 279

Qy 303 YVLNGTFSNLITVXAA-----APGP 322
Db 280 YLESGSVTAQVVLQAAIPLVSCGSPVPGTTDGYMPTAEAPGTTSRQGTTKTKVGGTTPGQ 339

Qy 323 CPPPPP-----PPRPS-----K 334
Db 340 MPTTQPSGTTVQMPTEVTATTSEQMLTSAVIDTTLAEVSTTEGTGTPTRPSGTTVAQ 399

Qy 335 PTPSLGPAGDNPL-----ELGRIPDE-----NCQINRYGHQFQATITVE 373
Db 400 ATTTEGPD-A-SPLPTQSSGTSISPLLDTTITMLVKRQVPLDCLVLYRGSFSLALDIVQ 458

Qy 374 GILEVNIITQMTDVLMPVPWPESSILDFVVTCCGSIPTVECTIISDPTCEITONTVCSPVD 433
Db 459 G-----IESAEILQAVPFSEGDAPFELTVSCQGLPKACMDISSPGCQPPAQRLCQSV 512

Qy 434 VDEMCLLTVRRTF-NGSGTYCVNLTGLDDTSLALTSTLISVPRDPASPLRMANSALISV 492
Db 513 PSPCQLVLHQVLKGGSGTYCLNVSADANSIAVASTQLVVFPGDGG-----LGOAPLLV 567
```

```
Qy 493 GCIAIFVTVI--SLLV----YKHKHEYN-----PIENSPGNVRSKGLSVFLNRAKAVFP 542
Db 568 GILLVAVVAVLASLILGDLRSRAQFPKCHMWALTAAPASGLRARGLG----- 615

Qy 543 GNQEKDPLLNQBF 555
Db 616 ---ENSPLLSQQ 625

RESULT 3
138400
melanoma-associated ME20 antigen (me20m) - human
N:Alternate names: melanoma antigen 25
C:Species: Homo sapiens (man)
C:Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 09-Jul-2004
C:Accession: I38400; A53668; A55753
R:Maresh, G.A.; Marken, J.S.; Neubauer, M.; Aruffo, A.; Hellstrom, I.; Hellstrom, K.; Ma;
DNA Cell Biol. 13, 87-95, 1994
A:Title: Cloning and expression of the gene for the Melanoma-Associated ME20 Antigen.
A:Reference number: I38400; MUID:94235165; PMID:8179825
A:Accession: I38400
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-662 <RES>
A:Cross-references: UNIPROT:P40967; UNIPARC:UPI000016A059; EMBL:U01874; NID:g494939; PID
R:Adema, G.J.; de Boer, A.J.; Vogel, A.M.; Loenen, W.A.M.; Figdor, C.G.
J. Biol. Chem. 269, 20126-20133, 1994
A:Title: Molecular characterization of the melanocyte lineage-specific antigen gp100.
A:Reference number: A53668; MUID:94327568; PMID:7519602
A:Accession: A53668
A:Molecule type: mRNA
A:Residues: 1-592,594-662 <ADE>
A:Cross-references: UNIPARC:UPI0000000C2D; GB:S73003; NID:g639589; PIDN:AACG0634.1; PID:
R:Kawakami, Y.; Eliyahu, S.; Delgado, C.H.; Robbins, P.F.; Sakaguchi, K.; Appella, E.; Y
Proc. Natl. Acad. Sci. U.S.A. 91, 6458-6462, 1994
A:Title: Identification of a human melanoma antigen recognized by tumor-infiltrating lym
A:Reference number: A55753; MUID:94294401; PMID:8022805
A:Accession: A55753
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-161,'P',163-592,594-662 <KAW>
A:Cross-references: UNIPARC:UPI000014ACC2
C:Keywords: glycoprotein

Query Match 20.6%; Score 622.5; DB 2; Length 662;
Best Local Similarity 25.5%; Pred. No. 1.8e-40;
Matches 168; Conservative 93; Mismatches 199; Indels 199; Gaps 17;

Qy 3 CLYYFLGFLLLAARLPDAAK--RFHDVLGNRPSPAYMREHNQNLGWSDDNDWNEKLYP 60
Db 8 CL----LHLAVIGALLAVGATKVPENQDLGVSR-----QLR-----TKAWNRLQYP 50

Qy 61 VWKRGDMWRKNSWKGRVQAVLTSDSPALVGSNTITFAVNLIPPRCKEDANGNIYVEKVC 120
Db 51 EWT--EAQLDCWRGGQVSLKVSNDGPTLVGNASFSIALNPFPGSQKVLPGQVIVWNT 108

Qy 121 RNEAGLSADPVVYN---W---TAMSEDSDCGENTGQSHHNVFPDGKPPFPHHPGRRWNF 173
Db 109 -----IINGSOVWGQGVPIPORTDDA-----CIFPDGPGPCSGSQKRSF 149

Qy 174 IYVFHTLQYFQKLRCSVRVSVNTANVTLPQLMEVTVYRRHG--RAYVPIAQKDVV 232
Db 150 VYVWKTGWQVQLGGPVSGLSIGTGRAMLGTHTMEVTVYHRRGSRSVYPLAHSSAFTI 209

Qy 233 TDQIPVFTVMPQKNDRSSDETFLKDLPIMPDVLIHDPSPHFLANTSTINYKWSFGDNTGLF 292
Db 210 TDQVPFVSVSQLRALDGGNKHFLRNQPLTFALQLHDPGVLAEADLSYTWDFGDSGTL 269

Qy 293 VSTNHTVNTHTVLTNGTFSNLITVXAAAP---GPCPPPP----- 327
Db 270 ISRALVTHTYLEPGPVTAQVVLQALPLTSCGSPVPGTTDGHRTAEAPNTTAGQVPT 329

Qy 328 ----- 328
```



## RESULT 6

A41105 protein-tyrosine-phosphatase (EC 3.1.3.48) PTPN4, nonreceptor type 4 [validated] - human  
N/Alternate names: PTPase MEG  
C/Species: Homo sapiens (man)  
C/Date: 20-Mar-1992 #sequence\_revision 02-May-1994 #text\_change 05-Oct-2004  
C/Accession: A41105  
R/Gu, M.; York, J.D.; Warshawsky, I.; Majerus, P.W.  
Proc. Natl. Acad. Sci. U.S.A. 88, 5867-5871, 1991  
A/Title: Identification, cloning, and expression of a cytosolic megakaryocyte protein-ty  
A/Reference number: A41105; MUID:91288564; PMID:1648233  
A/Molecule type: mRNA  
A/Accession: A41105  
A/Residues: 1-926 <GUA>  
A/Cross-references: UNIPROT:P29074; UNIPARC:UIP0000000CA9; GB:M68941; NID:g190747; PIDN:  
A/Experimental source: megakaryocytes, cell line MEG-10  
C/Genetics:

A/Gene: GDB:PTPN4  
A/Cross-references: GDB:I31387; OMIM:176878  
A/Map position: 9q31-9q31  
C/Superfamily: protein-tyrosine phosphatase, non-receptor types 3/4; GLGF domain homolog  
C/Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phosphatase  
F/31-308/Domain: protein 4.1 membrane-binding domain homology <B41>  
F/523-597/Domain: GLGF domain homology <GLG>  
F/679-900/Domain: protein-tyrosine-phosphatase homology <PTP>  
F/852/Active site: Cys (phosphocysteine intermediate) #status predicted  
F/858/Binding site: substrate phosphate (Arg) #status predicted

Query Match 4.5%; Score 136; DB 1; Length 926;  
Best Local Similarity 22.0%; Pred. No. 0.025;  
Matches 87; Conservative 43; Mismatches 125; Indels 140; Gaps 17;

```
Qy 40 EHNOLN-----GSSSDNDNWEKLYPWKGDMEKNSWKGRRVQAVLTSDSP 87
Db 204 EFNLYNTARTLELVGVBFHYARDQSN-NEIMIGVMSGILLIYKN-----RVRMNTFP 254

Qy 88 ALVGSNITFAVNLFFRCQEKDANG-----NTVYEKNCRNEAGLSADPYVYNWTAWSE 140
Db 255 WLKIVKISFKCKQFFIQLRKELHESRETLGLFNVMVYRACKN-----LWKACVE 303

Qy 141 DSDGENTGQSHHNVDPDGKPPHPHGWRRWNFIYVHTLIGQYFQKLGRCSV----- 192
Db 304 -----HHTFRLDRPLPPQK-----NFFAHYFTLGSKFRYCGRTEVQSVQYCKE 347

Qy 193 -----RVSVNTANVTGLGPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFVTMFQKND 247
Db 348 KANKDRVFARSPKPLARKLMDWEVSRN-----SISDRLETQSLRSRSPGTFNH 399

Qy 248 RNSS---DETLKDLPI--MFDVLIIH-DPSH-FLNYSTINYKSGFDNTGLFVSTNHTVN 300
Db 400 RNSFTQEGTGLRSPSSVGHLDVHMVHTSPSEVFVNQR-----SPSSTOA 443

Qy 301 HTYVLNGTFSNLNLTVKAAAPGCPPPPPPPR-----SKKNSWNQIHYGSHSQDLESHINETFDI 497
Db 444 NSIVLESSPS-----QETPDGKPPALPPRQSKKNSWNQIHYGSHSQDLESHINETFDI 497

Qy 332 ---PSKTPSLGPAGDNPLELSRIPDENCQINRYG 363
Db 498 PSSKEKTPNGGIPHDNLVLIRMKPDEN---GRFG 529
```

## RESULT 7

T07907 hydroxyproline-rich glycoprotein GAS28 precursor - Chlamydomonas reinhardtii  
C/Species: Chlamydomonas reinhardtii  
C/Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C/Accession: T07907  
R/Rodriguez-Martinez, H.; Haring, M.A.; Von Gromof, E.; Beck, C.F.  
submitted to the EMBL Data Library, July 1997  
A/Reference number: Z16207  
A/Accession: T07907  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA

A/Residues: 1-446 <ROD>  
A/Cross-references: UNIPROT:O22458; UNIPARC:UIP000000A8A0F; EMBL:AF015883; NID:g2384727;  
A/Experimental source: gametes  
C/Genetics:

A/Gene: GAS28

F/1-29/Domain: signal sequence #status predicted <SIG>  
F/30-446/Product: hydroxyproline-rich glycoprotein GAS28 #status predicted <MAT>

Query Match 4.0%; Score 120.5; DB 2; Length 446;  
Best Local Similarity 29.5%; Pred. No. 0.15; Mismatches 7; Indels 46; Gaps 6;  
Matches 38; Conservative 7

```
Qy 320 PGFCPPPPPPPPRPSKPTPSLGPA-----GDNPLELSRIPDENCQINRYGHFQ 366
Db 240 PPSPPPPPPPTTSPPPPELPAPQAPARKRPPPPASPPPSRDPFPCQQRNARGSRL 299

Qy 367 AT-----ITIVGILEVNI-IQMTDVLMPFVPWPESSLIDFVTCQGSIPTEVC----- 413
Db 300 MTTASNNITVVGILTRICFNVALKD-----CEN--PNSKCCFEFLYK 340

Qy 414 -TIISDPTC 421
Db 341 MEIETDPTC 349
```

## RESULT 8

C64483

hypothetical protein MJ1468 - Methanococcus jannaschii

C/Species: Methanococcus jannaschii

C/Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 09-Jul-2004

C/Accession: C64483

R/Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
son, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996

A/Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
A/Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
A/Reference number: A64300; MUID:96337999; PMID:8688087  
A/Accession: C64483

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1009 <BUL>

A/Cross-references: UNIPROT:Q58863; UNIPARC:UIP000013AB23; GB:U67587; GB:L77117; NID:g15  
C/Genetics:  
A/Map position: FOR1437031-1440060

Query Match 3.8%; Score 114.5; DB 2; Length 1009;  
Best Local Similarity 18.3%; Pred. No. 1.4;  
Matches 88; Conservative 58; Mismatches 157; Indels 177; Gaps 19;

```
Qy 100 LIIPRCQKEDANGNIVYEKNCRNEAGL-----SADPVVY----- 133
Db 490 LYPHVFIIDNGNI-----EVGIPPIGVGGYSYQIYASPIATPIYPIINITIV 541

Qy 134 ---NWT-AWSESDSGENGSTGQSHHNVFPDGKPPHPHGWRRWNFIYVHTLIGQYFQKLGR 189
Db 542 EPASWTWYV-----HHIYFGDGSVWIKPKKGPYTFHTVTSBGVY----- 582

Qy 190 CSVRVSVNTANVTGLGPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFVTMFQKNDRN 249
Db 583 -PIYMKVYTAE-----NMKTVVVIDNKNPIAKLYTYPNPAS 617

Qy 250 SSDS-----TFLKDLPIM-----FDVLIHDPSPHFLN 275
Db 618 YKDTISPSINSYDPPDANRMIPENYVYGTILIGYFIPSPNSPMARIYGFNLTYD----- 671

Qy 276 YSTINYKSGFDNTGLFVSTNHTV-NHT---YVLNGTFSNLNLTVKAAAPGCPPPP----- 326
Db 672 -SNGNVAWNYSSNELTIISKSPFIGNYTAFLVWDGSGVNSTTVKFSVINRPPVQAQFIY 730

Qy 327 -PPPPRPSKP-----TPSLGPAGDNPLELSRIPDENC-----QINRYGHFQATTTI 371
Db 731 YPDKPEPNEDVFEVSQSYDPEGEIAYIWNFGDGTIVNTTDTIVHHKYERPGYITVLTUV 790
```



Query Match 3.6%; Score 109.5; DB 2; Length 348;  
Best Local Similarity 21.2%; Pred. No. 0.76;  
Matches 83; Conservative 51; Mismatches 128; Indels 129; Gaps 22;  
QY 61 VKKGDWRKNSWG--GRVQAVLTSDSPALVGNITFAVNLIPRCKEDANGNIVYEK 118  
DB 16 VFETGQTKKLSATTGNDGTATAIIRN--AAGGGTVTVTCVL-----DADGKV--SD 63  
QY 119 NCR-NEAGLSADPVYVNTWMTASEDSGENGQSHHNVFDPGKPPHHGWRKRFVVF 177  
DB 64 QCDVHFAVASAD--IISFSVATDPLGLYSNNKGA-ARVFAEIIIPNNP---VTNCLINF 117  
QY 178 HTLG-----QYFOKLGRCSVRVSNVNTANVTGLPQLMEVTVRRHGRAYVPIAQVDVYV 231  
DB 118 STFGVSGIINNYVNIIEPN-----SVSSGSLMQ-----FVQIGNTKYT-- 157  
QY 232 VTDQIPVFTVMFQKNDNRNSDETFLKDLPIPMFDVLIHDPHFLNYSNTYKWSFGDNTGL 291  
DB 158 ----IPV-----DVSXKGPE-----EGSVGIQAAL 178  
QY 292 FVSTNHTVNHVYVLTNGFSL---NLTVKAAAGPCPPPPPPRPSKPTSLGPGADNPLE 348  
DB 179 VTTIQP-----DKAFALSTPNARVDFINPAP-PPPPPPSPCPPPPPPPPPPVVR 229  
QY 349 LS-----RIPDENCQINRYGHFQATITIVEGILEVNIQMTDVLMPVPWPBSS 396  
DB 230 PSYTIQIDAVVDNPIPTPGCRDIARVTVQNGRPLAGALVQCSLVSLG-----NTS 281  
QY 397 LIDFVVTCCGSIPTVCTIISDPCEITQNT 427  
DB 282 LM-PALAA-----NEIANLVSP-----QNT 301

## RESULT 12

Ti7414  
TipC protein - slime mold (Dictyostelium discoideum)  
C;Species: Dictyostelium discoideum  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T17414  
R;Steger, J.T.; Laub, M.T.; Loomis, W.F.  
submitted to the EMBL Data Library, July 1998  
A;Description: Interaction of tip genes in early Dictyostelium discoideum development.  
A;Reference number: Z18774  
A;Accession: T17414  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-3848 <STE>  
A;Cross-references: UNIPROT:O76737; UNIPARC:UPI0000076F0B; EMBL:AF079445; NID:G3420744;  
A;Experimental source: strain AX4  
C;Genetics:  
A;Gene: tipC  
A;Introns: 72/3  
C;Function:  
A;Description: required for tip formation

Query Match 3.6%; Score 108; DB 2; Length 3848;  
Best Local Similarity 19.0%; Pred. No. 28;  
Matches 69; Conservative 55; Mismatches 132; Indels 108; Gaps 14;  
QY 3 CLYYFLGFLLAARLPDAKRFDHVLGNERPSAYMREHNLGNSDENDWNEKLYPVW 62  
DB 918 QQYVNLNIIQSVTPIDQQQ--LIGSE--SMYSNADFSSVYQSNIFKQPQLOLQ 972  
QY 63 KRGMWRKNSWKGROVALTSDSPALVGSNITFAVNLIPRCKEDANGNIVYEKCRN 122  
DB 973 KQ-----QSPPLTSSP-----PLIKKKKSYSEKFDISYKMF 1003  
QY 123 EAGLSADPY-----VNTWTAMSEDSGENTGQSH-----HNVFDPGKPPHHPGWR 170  
DB 1004 EASNIDKFNHLRLDNDYSSGSSNNDSSNTTTNNQLVSHV----- 1048  
QY 171 WNFIVFHTLGGYFQKLGRCRVRVSVN---TANVTGLPQLMEVTVRRHGRAYVPIAQVK 227

DB 1049 -ALVYLRCLHGTYVQKFLFDTNLELLIKGMWIEDCFQKPSI-----AHGSG----- 1093  
QY 228 DVYVVTQDIPFVFTMFQKNDRN-----SSDETFLKDLPIPMEDVLIHDP 270  
DB 1094 -----GDLATTTNRFPKPDVNFVGGSNENLIKFKIQIQQISSDSPFYNNIDKLLDIELSQI 1148  
QY 271 SHFLANTSTYKWSFGDN-TGLFVSTNHTVNHVTVVLTNGTFSNLNLTVKAAPG---PCPPP 326  
DB 1149 NLILNRKTVAGLIEFSSVSNLSILKXNNQNNQ---NNNQNNQNNINESFTVITSPPP 1205  
QY 327 PPPP 330  
DB 1206 PPPP 1209

## RESULT 13

T23694  
hypothetical protein M03C11.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T23694  
R;McMurray, A.  
submitted to the EMBL Data Library, April 1995  
A;Reference number: Z19783  
A;Accession: T23694  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-848 <WIL>  
A;Cross-references: UNIPROT:Q21489; UNIPARC:UPI0000079560; EMBL:Z49128; PIDN:CAA88959.1;  
A;Experimental source: clone M03C11  
C;Genetics:  
A;Gene: CESP-M03C11.2  
A;Map position: 3  
A;Introns: 113/2; 147/3; 185/3; 379/1; 482/3; 553/2; 688/3; 762/3

Query Match 3.5%; Score 106.5; DB 2; Length 848;  
Best Local Similarity 20.8%; Pred. No. 4.5;  
Matches 117; Conservative 75; Mismatches 203; Indels 167; Gaps 30;  
QY 83 TSDSPALV--GSNITFAVN-----LIIFPRCKEDANGNIVYEKCRNEAGLSADP 130  
DB 206 TRFQPRIVTCSAGRTLCVNEEVKKLKLNHLINELKMLRKNKGSEKEKVQKLEKGTTKT 265  
QY 131 YV-----VNTWTAMSEDSG-----ENGTGOSHNVFPDGKPF----- 162  
DB 266 KTCATSCFYNSTQIEDVWNGVLSNKLSTLEVSQKGLNSGCPYFATRKSVFQCVLL 325  
QY 163 PH-----HPGWRR-W-----NFTY-----VFHTLGOYF-----QKLGK-----CSV-RVS 195  
DB 326 PYQVLLHDGTRKAWGIELKDNVILDEAHNVLTISFFFRKLEKSVKNGFKNCLSYAE 385  
QY 196 VNTANVTGLPQLMEVTVRRHGRAYV---PIAQVDVYVVTQDIPFVFTMFQKNDNRSSD 252  
DB 386 ISTKSLTLAURL--IREYNAHYKLLLAHNLNLYKQLESILTSMKLIFLNSQSKEDVMTMA 443  
QY 253 ETEKDLPIPMFDVLIHDPHFLNYSNTYKWSFGDNTGLFVSTN-----HTVNHVYVLTNGT 308  
DB 444 Q-LARNILNII-EINLPKLAEMYEMKTDLCCKFH-----GFYMLQKEIKKENKPKLTGI 496  
QY 309 FSNLTVKAAAPGCPPPPPPPRPSKPTP-----SLGPAGDNPLELSRIIDENCQINRYG 363  
DB 497 QKL-MAAKEAPEPEAEPLPPKP-VPSPLFSLSKSFIDALTNKCEDGRIIIVEKSATEAKF 554  
QY 364 HF-----QATITIVEGILEVNIQMTDVLMPVPWPBSSLIDFVFTVTCQS 407  
DB 555 RFMLLNADRLSEVWTSARATI-LVGGTME-----PAQLLVE--TLSRGS 596  
QY 408 IPTVEVCTIISDPCEITQNTVCSFVDVDEMCLLTVTRTFNGSGTYCVNLTLGDDTSALT 467  
DB 597 IGAD-----SIRRFSCCHVIDDQLAVTVERTVDGKPFQLTQYTRGADTTLSRL 646  
QY 468 STLISVPDRDPASPRLRMANSLIS--VGCLAI FVTWISLL--VYKHKHENPIENSPGNV 523

Db 647 ATSI-----QALPHDPNGVVFVPSYDFLNFQKKKEF----- 681  
QY 524 VRSGKLSVFLNRAKAVFFPQNG 545  
Db 682 ----GILKRIEKKAVFTESRQ 699

## RESULT 14

JQ0533  
genome polyprotein - Kennedy yellow mosaic virus (strain Jervie Bay)  
N/Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase  
N/Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C/Species: Kennedy yellow mosaic virus  
C/Date: 17-Feb-1994 #sequence revision 17-Feb-1994 #text change 09-Jul-2004  
C/Accession: JQ0533  
R/Ding, S.; Keese, P.; Gibbs, A.  
J. Gen. Virol. 71, 925-931, 1990  
A/Title: The nucleotide sequence of the genomic RNA of Kennedy yellow mosaic tymovirus-  
A/Reference number: JQ0533; MUID:90218040; PMID:2324710  
A/Accession: JQ0533  
A/Molecule type: genomic RNA  
A/Residues: 1-1874 <DIN>  
A/Cross-references: UNIPROT:P36304; UNIPARC:UPI0000131EA7; GB:D00637; NID:9221969; PIDN:  
C/Superfamily: eggplant mosaic virus RNA-directed RNA polymerase  
C/Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; R  
F:1002-1009/Region: nucleotide-binding motif A (P-loop)  
F:1064-1069/Region: nucleotide-binding motif B  
F:1008/Binding site: ATP (lys) #status predicted

Query Match 3.5%; Score 106; DB 1; Length 1874;  
Best Local Similarity 20.2%; Pred. No. 15;  
Matches 75; Conservative 38; Mismatches 100; Indels 158; Gaps 16;

QY 257 KDLPIMFVLIHDPSHFLNYSTIN-----YKWSFGDNTGLFVS 294  
Db 526 KDIPLL-----PR--MSFTVNLPLQPPMMLAIGASLVPELFLSWLSGD---VDLQ 573  
QY 295 TNHTVNTVYVLTNGTSLNLT---VKAAAGP-----CP--PPPPPPRPS 333  
Db 574 TQHDIVYHHLHPENFTLSWTRPYLALADSPFLVAHSPPLPVNSSLPPPPPPPLPES 633  
QY 334 KTPSLGP-----AGNPLELSRIPDE 355  
Db 634 QPLSQGPATQAPSAQPTGCEPLAPPTTELKPSNPNPNPSSAGSNPPPKSSSDN 693  
QY 356 NCQINRYGHFOATIT-----IVGILEVNI-----QMTDVLMPVP 391  
Db 694 PAPNKPPTSSSTPPSNPLQFGSIHSPFLSQNLNYSALPPQDPTNTLSLPEP 753  
QY 392 WPESSLIDFVVTQGSIPTEV-CTIISDPTCBITQVCSYVDVDEMCLLTVRRTFNGSG 450  
Db 754 KP-----PTEVQSPLMADPTC-----VGPAVSFSS---LYPRDFFPNTA 789  
QY 451 TVCVNLTGDDTSIALTST---LISVDRDPSAPLR-----MANSALISVGC 494  
Db 790 SPLTFLRLSPPTPLPMPKNNCLLTAVAPSLHNPRLMTSLQEVLPDLSLSNDSVGM 849  
QY 495 LAIFVTVISLL 505  
Db 850 STDILTALSHL 860

## RESULT 15

C96521  
protein F21D18.18 [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence revision 02-Mar-2001 #text change 09-Jul-2004  
C/Accession: C96521  
R/Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: C96521  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-492 <STO>  
A/Cross-references: UNIPROT:O9LNG3; UNIPARC:UPI00000AA2C0; GB:AE005173; NID:98778527; P:  
C/Genetics:  
A/Genes: F21D18.18  
A/Map position: 1

Query Match 3.5%; Score 105; DB 2; Length 492;  
Best Local Similarity 24.7%; Pred. No. 2.8;  
Matches 41; Conservative 10; Mismatches 43; Indels 72; Gaps 8;  
QY 272 HFLNYSTINYKWSFGDNTGLFVSTNHTVNTVYVLTNGTSLNLTVKAAAGPCCPPPPPPPPR 331  
Db 20 HFTNVARARYHYHG-----RHGVTH-----PLPPPPPPPL 50  
QY 332 -----PSKPTPSLQPA-GDNP-----LELSRIPDENCQ----- 358  
Db 51 ETANPPDQVPSPDYFSPDPAPGDSGCVDFVTSTFGAVGDSGDDTAAPQDAWKAACAVE 110  
QY 359 -----INRYGHFOATITVEGILEVNI-----QMTDVLN-----PVPWPE 394  
Db 111 SGVLAPEGGVFKITITIFSGCKGLVFLQDGLVMPDPGPEWPE 156

Search completed: December 24, 2005, 04:36:31

Job time : 48 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:25:06 ; Search time 233 Seconds  
(without alignments)  
1695.690 Million cell updates/sec

Title: US-09-853-880A-17  
Perfect score: 3026  
Sequence: 1 MECUYFLGILLARLPD.....PFGNQKDPKLNQBFKGV 560

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	3010	99.5	572 1	GNPMB_HUMAN
2	2197	72.6	574 1	GNPMB_MOUSE
3	2179	72.0	572 1	GNPMB_RAT
4	2163	71.5	572 2	Q9QZF6_RAT
5	2042	67.5	526 2	Q8BVA0_MOUSE
6	1588.5	52.5	559 1	QNR71_COTJA
7	1431.5	47.3	595 2	Q4V805_XENLA
8	1029	34.0	206 2	Q96F58_HUMAN
9	1022	33.8	206 2	Q8IXJ5_HUMAN
10	843	27.9	538 2	Q4RKH9_TETNG
11	738.5	24.4	719 2	Q4KLIV9_XENLA
12	735.5	24.3	721 2	Q6DDN6_XENLA
13	729.5	24.1	746 2	Q6DIR2_XENLA
14	682.5	22.6	808 2	Q52KZ4_XENLA
15	671.5	22.2	626 2	Q9CZB2_MOUSE
16	667.5	22.1	845 2	Q4VW61_BRARE
17	663.5	21.9	636 2	Q4LE84_PIG
18	659	21.8	626 1	PME17_MOUSE
19	652	21.5	604 2	Q4RIM0_TETNG
20	651	21.5	733 2	Q93391_COTCO
21	650	21.5	763 1	PME17_CHICK
22	649.5	21.5	703 2	Q4R7T2_MACFA
23	622	20.6	661 1	PME17_HUMAN
24	561	18.5	413 2	Q4S699_TETNG
25	459	15.2	491 1	PME17_BOVIN
26	442.5	14.6	251 2	Q5R2V2_BRARE
27	399	13.2	461 2	Q97884_HORSE
28	247.5	8.2	236 2	Q9QY67_MOUSE
29	182.5	6.0	423 2	Q5R6F5_PONPY
30	182	6.0	397 2	Q8N3R2_HUMAN
31	181.5	6.0	423 2	Q8N0W9_HUMAN

32	181	6.0	435	2	Q8N3G9_HUMAN	Q8n399 homo sapien
33	167	5.5	141	2	Q9QY70_MOUSE	Q9qy70 mus musculus
34	156	5.2	419	2	Q6NXX3_MOUSE	Q6nxx3 mus musculus
35	155.5	5.1	3779	2	Q4RZY1_TETNG	Q4rzy1 tetraodon n
36	153	5.1	321	2	Q8IY46_HUMAN	Q8iy46 homo sapien
37	144	4.8	1988	2	Q8TPZ1_METAC	Q8tpz1 methanosarc
38	142.5	4.7	1817	2	Q8TI59_METAC	Q8ti59 methanosarc
39	140	4.6	688	2	Q8TR88_METAC	Q8tr88 methanosarc
40	137	4.5	906	2	Q8TPY9_METAC	Q8tpy9 methanosarc
41	136	4.5	926	1	PTN4_HUMAN	P29074 homo sapien
42	133	4.4	881	2	Q6H7U3_ORYSA	Q6h7u3 oryza sativ
43	132	4.4	331	2	Q4V9B2_BRARE	Q4v9b2 brachydanio
44	127.5	4.2	603	2	Q7U5X8_SYNPK	Q7u5x8 synecococc
45	126.5	4.2	514	2	Q7PYC8_ANOGA	Q7pyc8 anopheles g

ALIGNMENTS

RESULT 1  
GNPMB\_HUMAN  
ID GPNMB\_HUMAN STANDARD; PRT; 572 AA.  
AC Q14956; Q6UVX1; Q8N1A1;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 10-MAY-2005 (Rel. 47, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE Transmembrane glycoprotein NMB precursor (Transmembrane glycoprotein HGFN).  
GN Name=GNPMB; Synonyms=HGFN, NMB; ORFNames=UNQ1725/PRO9925;  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).  
RC TISSUE=Melanoma;  
RX MEDLINE=95113576; PubMed=7814155;  
RA Waterman M.A.J., Ajubi N., van Dinter I.M.R., Degen W.G.J.,  
RA van Muljen G.N.P., Ruiter D.J., Bloemers H.P.J.;  
RT "NMB, a novel gene, is expressed in low-metastatic human melanoma cell lines and xenografts.";  
RL Int. J. Cancer 60:73-81 (1995).  
RN [2]  
RP NUCLEOTIDE SEQUENCE (ISOFORM 2).  
RC TISSUE=Peripheral blood;  
RX MEDLINE=22498106; PubMed=12609765; DOI=10.1016/S0167-0115(02)00288-4;  
RA Bandari P.S., Qian J., Yehia G., Joshi D.D., Maloof P.B., Potian J.,  
RA Oh H.S., Gascon P., Harrison J.S., Rameshwar P.;  
RT "Hematopoietic growth factor inducible neurokinin-1 type: a transmembrane protein that is similar to neurokinin 1 interacts with substance P.";  
RL Regul. Pept. 111:169-178 (2003).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,  
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.R., Heldens S.,  
RA Huang A., Kim H.S., Klinowski L., Jin Y., Johnson S., Lee J.,  
RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,  
RA Seahagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
RA Vandlen R.L., Watanabe C., Wieand D., Woods K., Xie M.-H.,  
RA Yansura D.G., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,  
RA Wood W.I., Godowski P.J., Gray A.M.;  
RT "The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment.";  
RL Genome Res. 13:2265-2270 (2003).  
RN [4]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RX MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;  
RA Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,

RA Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,  
 RA Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,  
 RA Fowell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,  
 RA Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,  
 RA Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,  
 RA Ozerkany P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,  
 RA Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,  
 RA Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,  
 RA Tin-Wollan A.-M., Abbott A., Minx P., Maupin R., Strowmatt C.,  
 RA Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,  
 RA Wendt M.C., Yang S.-P., Schultz B.R., Wallis J.W., Spieth J.,  
 RA Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.B.,  
 RA Hertenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,  
 RA Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,  
 RA Gillett W., Zhou Y., James R., Phelps K., Iadonoto S., Bubbs K.,  
 RA Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Pury T.S.,  
 RA Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,  
 RA Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,  
 RA Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,  
 RA Waterston R.H., Wilson R.K.;  
 RT "The DNA sequence of human chromosome 7."; ;  
 RL Nature 424:157-164 (2003). ;  
 RN [5]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J.G., Heide F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grinwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -!- FUNCTION: Could be a melanogenic enzyme (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q14956-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q14956-2; Sequence=VSP\_013001;  
 CC -!- TISSUE SPECIFICITY: Not restricted to the melanocytic lineage.  
 CC -!- DEVELOPMENTAL STAGE: Expression in poorly metastatic melanoma cell  
 CC lines; no expression in highly metastatic melanoma cell lines.  
 CC -!- SIMILARITY: Belongs to the Pmel-17/NMB family.  
 CC -!- SIMILARITY: Contains 1 PKD domain.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use as long as its content is in no way modified and this statement is not  
 CC removed.  
 CC -----  
 CC EMBL; X76534; CRA54044.1; -; mRNA.  
 CC EMBL; AF322909; AAG42839.1; -; mRNA.  
 CC EMBL; AY359124; AAO89481.1; -; mRNA.  
 CC EMBL; AC005082; AAP22336.1; -; Genomic\_DNA.  
 CC EMBL; BC032783; AAH32783.1; -; mRNA.  
 CC F01; I38065; I38065.  
 CC EMBL; ENSG00000136235; Homo sapiens.

DR HGNC; HGNC:4462; GPNMB.  
 DR MIM; 604368; -.  
 DR GO; GO:0016021; C:integral to membrane; TAS.  
 DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR000601; PKD.  
 DR SMART; SM00089; PKD; 1.  
 DR PROSITE; PS50093; PKD; 1.  
 KW Alternative splicing; Glycoprotein; Polymorphism; Signal;  
 KW Transmembrane.  
 FT SIGNAL 1 21 Potential.  
 FT CHAIN 22 572 Transmembrane glycoprotein NMB.  
 FT TOPO\_DOM 22 486 Extracellular (Potential).  
 FT TRANSMEM 487 507 Potential.  
 FT TOPO\_DOM 508 572 Cytoplasmic (Potential).  
 FT DOMAIN 240 327 PKD.  
 FT MOTIF 64 66 Cell attachment site (Potential).  
 FT COMPTAS 320 332 Poly-Pro. (Potential).  
 FT CARBOHYD 93 93 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 134 134 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 146 146 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 200 200 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 249 249 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 275 275 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 300 300 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 306 306 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 312 312 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 459 459 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 467 467 N-linked (GlcNAc...) (Potential).  
 FT VARSPPLIC 340 351 Missing (in isoform 2).  
 FT VARIANT 195 195 S -> C (in dbSNP:530436).  
 FT VARIANT 197 197 N -> H (in dbSNP:530413).  
 FT CONFLICT 354 354 A -> T (in Ref. 3).  
 FT SEQUENCE 572 AA; 63923 MW; 2465C12CF0F3996 CRC64;  
 Query Match 99.5%; Score 3010; DB 1; Length 572;  
 Best Local Similarity 97.9%; Pred. No. 1.le-218;  
 Matches 560; Conservative 0; Mismatches 0; Indels 12; Gaps 1;  
 Qy 1 MECLYPLGLFLLAARLPDAARFHDVLGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60  
 Db 1 MECLYPLGLFLLAARLPDAARFHDVLGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60  
 Qy 61 VVKRGDMRWKNSKGGVQAVLTSDSPALVGSNTTFAVNLIFPCQKEDANGNIYKNC 120  
 Db 61 VVKRGDMRWKNSKGGVQAVLTSDSPALVGSNTTFAVNLIFPCQKEDANGNIYKNC 120  
 Qy 121 RNEAGLSADPVYVNTWTAWSDSDGNGTQSHNVFPDGPFPHPHGRWRNFIYVFTLL 180  
 Db 121 RNEAGLSADPVYVNTWTAWSDSDGNGTQSHNVFPDGPFPHPHGRWRNFIYVFTLL 180  
 Qy 181 GQYFQKLGRCSRVSVNTANTVTGQPMVTVVRRHGRVYVIAQVXDVVVTVQIPVFV 240  
 Db 181 GQYFQKLGRCSRVSVNTANTVTGQPMVTVVRRHGRVYVIAQVXDVVVTVQIPVFV 240  
 Qy 241 TFMQKDRNSDSETFLLKDLPIFMEDVLIHDPHSFLNYSTINYKSPGNTGLFVSTNHTVN 300  
 Db 241 TFMQKDRNSDSETFLLKDLPIFMEDVLIHDPHSFLNYSTINYKSPGNTGLFVSTNHTVN 300  
 Qy 301 HTVYVNGTFFSLNLTVAKAAPGCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 348  
 Db 301 HTVYVNGTFFSLNLTVAKAAPGCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 360  
 Qy 349 LSRIPDENCQINRGYHGFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVTCQSSI 408  
 Db 361 LSRIPDENCQINRGYHGFQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVTCQSSI 420  
 Qy 409 PTEVCTIISDPTCEITONTVCSPVDVEMCLLTVRRRTFNGSGTYCNVLTGDDTSLAITS 468  
 Db 421 PTEVCTIISDPTCEITONTVCSPVDVEMCLLTVRRRTFNGSGTYCNVLTGDDTSLAITS 480

QY 469 TLISVDPDRPASPPIKMANSLISVCLAFVTVISLLVYKKHKNYPIENSPGNVVRSGK 528  
 DB 481 TLISVDPDRPASPPIKMANSLISVCLAFVTVISLLVYKKHKNYPIENSPGNVVRSGK 540

QY 529 LSVFLNRAKAVFPFGNQEKDPLLNQKQEPKGV 560  
 DB 541 LSVFLNRAKAVFPFGNQEKDPLLNQKQEPKGV 572

RESULT 2

GNMB MOUSE  
 ID GPNMB MOUSE STANDARD; PRT; 574 AA.  
 AC Q99P91; Q8BV9; Q8XL4; Q9QXA0;  
 DT 10-MAY-2005 (Rel. 47, Created)  
 DT 10-MAY-2005 (Rel. 47, Last sequence update)  
 DT 10-MAY-2005 (Rel. 47, Last annotation update)  
 DE Transmembrane glycoprotein NMB precursor (Dendritic cell-associated transmembrane protein) (DC-HIL).  
 GN Name=Gpmb; Synonyms=Dchil, Nmb;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=BA1B/C;  
 RX MEDLINE=21269372; PubMed=11114299; DOI=10.1074/jbc.M008539200;  
 RA Shikano S., Bonkobata M., Zukas P.K., Arizumi K.;  
 RT "Molecular cloning of a dendritic cell-associated transmembrane protein, DC-HIL, that promotes RGD-dependent adhesion of endothelial cells through recognition of heparan sulfate proteoglycans.";  
 RL J. Biol. Chem. 276:8125-8134 (2001).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE, TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.  
 RX PubMed=12638126; DOI=10.1016/S1567-133X(02)00012-1;  
 RA Bachner D., Schroder D., Grosse G.;  
 RT "mRNA expression of the murine glycoprotein (transmembrane) nmb (Gpmb) gene is linked to the developing retinal pigment epithelium and iris.";  
 RL Gene Expr. Patterns 1:159-165 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=C57BL/6J; TISSUE=Retina, and Skin;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Sult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmer S., Gusninch S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKensie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Iehi Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].  
 RC STRAIN=FVB/N; TISSUE=Salivary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 CC -1- FUNCTION: Could be a melanogenic enzyme (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Vesicular, endosomal like structures.  
 CC -1- TISSUE SPECIFICITY: In early development mRNA is detected at high levels in the outer layer of the retina. Later in development expression gets restricted to the retinal pigment epithelium and iris.  
 CC -1- SIMILARITY: Belongs to the Pmel-17/NMB family.  
 CC -1- SIMILARITY: Contains 1 PKD domain.  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.  
 CC EMBL; AF322054; AAK14240.1; -; mRNA.  
 CC EMBL; AJ251685; CAB85272.1; -; mRNA.  
 CC EMBL; AK044764; BAC32074.1; -; mRNA.  
 CC EMBL; AK076347; BAC36306.1; -; mRNA.  
 CC EMBL; BC026375; AAH26375.1; -; mRNA.  
 CC Ensembl; ENSMUSG0000029816; Mus musculus.  
 CC MGI; MGI:1934765; Gpmb.  
 CC GO; GO:0005615; C:extracellular space; TAS.  
 CC GO; GO:0005887; C:integral to plasma membrane; IDA.  
 CC GO; GO:0008201; F:heparin binding; IDA.  
 CC GO; GO:0005178; F:integrin binding; IDA.  
 CC GO; GO:0007155; P:cell adhesion; IDA.  
 CC InterPro; IPR000601; PKD.  
 CC SMART; SM00089; PKD; 1.  
 CC PROSITE; PS00093; PKD; 1.  
 KW Glycoprotein; Signal; Transmembrane.  
 FT SIGNAL 1 22 Potential.  
 FT CHAIN 23 574 Transmembrane glycoprotein NMB.  
 FT TOFO DOM 23 502 Extracellular (Potential).  
 FT TRANS MEM 503 523 Potential.  
 FT TOFO DOM 524 574 Cytoplasmic (Potential).  
 FT DOMAIN 250 338 PKD.  
 FT MOTIF 558 560 Cell attachment site (Potential).  
 FT CARBOHYD 93 93 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 134 134 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 200 200 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 249 249 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 275 275 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 296 296 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 300 300 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 306 306 N-linked (GlcNAc...) (Potential).

```
FT CARBOHYD 312 312 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 463 463 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 471 471 N-linked (GlcNAc... ) (Potential).
FT CONFLICT 36 36 N -> D (in Ref. 2 and 3; BAC36306).
FT CONFLICT 333 333 S -> S (in Ref. 2).
FT CONFLICT 339 341 SPP -> LPS (in Ref. 2).
FT CONFLICT 498 498 L -> M (in Ref. 3; BAC32074).
FT CONFLICT 541 541 K -> E (in Ref. 3; BAC32074).
SQ SEQUENCE 574 AA; 63675 MW; 6A8E9AC27AE6ACD0 CRC64;

Query Match 72.6%; Score 2197; DB 1; Length 574;
Best Local Similarity 70.7%; Pred. No. 2.9e-157;
Matches 403; Conservative 69; Mismatches 82; Indels 16; Gaps 2;

Qy 1 MECIYFLGLLLAARLPDAARFHDVILGNRPSPAYMRHNLNGWSSDENWEXLYP 60
Db 1 MESLCGVLLAAGLPQAARFHDVILGNRPSPAYMRHNLNGWSSDENWEXLYP 60

Qy 61 VWRGDMRWKNSWGGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VWRGDMRWKNSWGGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNWNTAWSEDSGNGTQSHHNVPPDGKPPHHPGRWNNFYVFTL 180
Db 121 RNDLGSLDLVYNWNTAGDGDWEDGTSRQHLRFPDRRPPRPHGKMSFYVFTL 180

Qy 181 GOYFQKLGRCVRSVNTANVTLPOLMEVTVYRRHGRVYPIAQKDVVVVDIPIVFV 240
Db 181 GOYFQKLGRCVRSVNTANVTLPOLMEVTVYRRHGRVYPIAQKDVVVVDIPIVFV 240

Qy 241 TMFQKNDNRSSDETFLLKDLPIFMFDVLIHDPSPHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMSQKNDNRSSDETFLLKDLPIFMFDVLIHDPSPHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPDAGD 344
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPDAGD 344

Qy 361 KSMELSDISNENCRINRYGFRAITIVGILEVSIQIADVPPTOPANSLMDFVTC 420
Db 361 KSMELSDISNENCRINRYGFRAITIVGILEVSIQIADVPPTOPANSLMDFVTC 420

Qy 405 QGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 464
Db 405 QGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 464

Qy 421 KGAFPEACTIISDPTCQIAQNRVCSFVAVDGLCLLSVRAFNAGSGTYCVNFTLGDDASL 480
Db 421 KGAFPEACTIISDPTCQIAQNRVCSFVAVDGLCLLSVRAFNAGSGTYCVNFTLGDDASL 480

Qy 465 ALTSTLISVDRDPASPLRMANSALISVGCIAIFVTVISLIVYKHKKEYNPIENSPGNV 524
Db 465 ALTSTLISVDRDPASPLRMANSALISVGCIAIFVTVISLIVYKHKKEYNPIENSPGNV 524

Qy 525 RSKGLSVPLNRAKAVFFPGNQEKDPLLNQ 554
Db 525 RSKGLSVPLNRAKAVFFPGNQEKDPLLNQ 554

Qy 541 KKGKGLSVLLSHAKAPFFRGDQEKDPLLDQK 570
Db 541 KKGKGLSVLLSHAKAPFFRGDQEKDPLLDQK 570
```

## RESULT 3

```
GNMB RAT
ID QPNMB RAT STANDARD; PRT; 572 AA.
AC Q6P7C7;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Transmembrane glycoprotein NMB precursor.
GN Name=Gnmb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Prostate;
RG NIH - Mammalian Gene Collection (MGC) project;
```

```
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Could be a melanogenic enzyme (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the Pmel-17/NMB family.
CC -!- SIMILARITY: Contains 1 PKD domain.
CC -----
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; BC061725; AAH61725.1; -; mRNA.
DR Ensembl; ENSRNOG0000008816; Rattus norvegicus.
DR RGD; 71008; Gnmb.
DR InterPro; IPR000601; PKD.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00093; PKD; 1.
KW Glycoprotein; Signal; Transmembrane.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 572 Transmembrane glycoprotein NMB.
FT TOPO_DOM 23 500 Extracellular (Potential).
FT TRANSMEM 501 521 Potential.
FT TOPO_DOM 522 572 PKD.
FT DOMAIN 251 338 PKD.
FT MOTIF 556 558 Cell attachment site (Potential).
FT CARBOHYD 93 93 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 134 134 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 200 200 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 249 249 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 275 275 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 296 296 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 300 300 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 306 306 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 312 312 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 461 461 N-linked (GlcNAc... ) (Potential).
FT CARBOHYD 469 469 N-linked (GlcNAc... ) (Potential).
SQ SEQUENCE 572 AA; 63731 MW; 99854F7773FF946C CRC64;

Query Match 72.0%; Score 2179; DB 1; Length 572;
Best Local Similarity 69.7%; Pred. No. 6.6e-156;
Matches 396; Conservative 74; Mismatches 84; Indels 14; Gaps 1;

Qy 1 MECIYFLGLLLAARLPDAARFHDVILGNRPSPAYMRHNLNGWSSDENWEXLYP 60
Db 1 MESLCGVLLAAGLPQAARFHDVILGNRPSPAYMRHNLNGWSSDENWEXLYP 60

Qy 61 VWRGDMRWKNSWGGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
Db 61 VWRGDMRWKNSWGGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNWNTAWSEDSGNGTQSHHNVPPDGKPPHHPGRWNNFYVFTL 180
Db 121 RNEAGLSADPYVYNWNTAWSEDSGNGTQSHHNVPPDGKPPHHPGRWNNFYVFTL 180

Qy 181 GOYFQKLGRCVRSVNTANVTLPOLMEVTVYRRHGRVYPIAQKDVVVVDIPIVFV 240
Db 181 GOYFQKLGRCVRSVNTANVTLPOLMEVTVYRRHGRVYPIAQKDVVVVDIPIVFV 240

Qy 241 TMFQKNDNRSSDETFLLKDLPIFMFDVLIHDPSPHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRSSDETFLLKDLPIFMFDVLIHDPSPHFLNYSITNYKWSFGDNTGLFVSTNHTVN 300

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPDAGD 346
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKTPSLGPDAGD 346

Qy 361 KSMELSDISNENCRINRYGFRAITIVGILEVSIQIADVPPTOPANSLMDFVTC 360
Db 361 KSMELSDISNENCRINRYGFRAITIVGILEVSIQIADVPPTOPANSLMDFVTC 360

Qy 405 QGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 464
Db 405 QGSIPTEVCTIISDPTCEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSL 464

Qy 421 KGAFPEACTIISDPTCQIAQNRVCSFVAVDGLCLLSVRAFNAGSGTYCVNFTLGDDASL 480
Db 421 KGAFPEACTIISDPTCQIAQNRVCSFVAVDGLCLLSVRAFNAGSGTYCVNFTLGDDASL 480

Qy 465 ALTSTLISVDRDPASPLRMANSALISVGCIAIFVTVISLIVYKHKKEYNPIENSPGNV 524
Db 465 ALTSTLISVDRDPASPLRMANSALISVGCIAIFVTVISLIVYKHKKEYNPIENSPGNV 524

Qy 525 RSKGLSVPLNRAKAVFFPGNQEKDPLLNQ 554
Db 525 RSKGLSVPLNRAKAVFFPGNQEKDPLLNQ 554

Qy 541 KKGKGLSVLLSHAKAPFFRGDQEKDPLLDQK 570
Db 541 KKGKGLSVLLSHAKAPFFRGDQEKDPLLDQK 570
```

```

Db 421 ATPTACTIISDPTCQIAQNRVCSFVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 480
Qy 467 TSTLSVPRDRPASPLRMANSALISVGCLAFVTVISLLVYKHKYKYNPIENSPGNVRS 526
Db 481 TSALISIPGKDLGSLRTVNGVLISIGCLAFVTVISLLVYKHKYKYNPIENSPGNVRS 540
Qy 527 KGLSVFLNRAKAVFPFGNQEKDPLLNQ 554
Db 541 KGLSVFLSHAKAFPSRGDREKDPQLQDK 568

RESULT 4
OSQZF6 RAT PRELIMINARY; PRT; 572 AA.
AC OSQZF6 RAT PRELIMINARY; PRT; 572 AA.
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Osteoactivin.
GN Name=Gpmb;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21611396; PubMed=11746512; DOI=10.1002/jcb.1259;
RA Safadi F.F., Xu J., Smock S.L., Rico M.C., Owen T.A., Popoff S.N.;
RT "Cloning and characterization of osteoactivin, a novel cDNA expressed
in osteoblasts.";
RL J. Cell. Biochem. 84:12-26(2001).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Odgren P.R., Marks S.C. Jr.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Xu J., Safadi F.F., Rosenzweig A.B., Popoff S.N.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184983; AAF03400.1; -; mRNA.
DR Ensembl; ENSRNOG0000008816; Rattus norvegicus.
DR RGD; 71008; Gpmb.
DR InterPro; IPR000601; PKD.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS0093; PKD; 1.
SQ SEQUENCE 572 AA; 63772 MW; CB939FB043EC57EB CRC64;

Query Match 71.5%; Score 2163; DB 2; Length 572;
Best Local Similarity 69.4%; Pred. No. 1.1e-154;
Matches 394; Conservative 75; Mismatches 85; Indels 14; Gaps 1;

Qy 1 MECLYVFLGLLAAALPLDAKRFHVLGNRPSPAYMREHNLGWSDDENNEKLYP 60
Db 1 MESLCGVFLVLLAAGLPQAARFRDVLGHEQYDPDHRENNQLRGWSDDENNEKLYP 60
Qy 61 VVKRGDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYERNC 120
Db 61 VVRGEGRWKDSWEGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYERNC 120
Qy 121 RNEAGLSADPYVYNNTANSESDGNGTQSQSHNVNPPDGKPPHPHGRWRNFIYVFTL 180
Db 121 RSDLEASDPYVYNNTGADDEWEDNTSQGHLPDPGKPPRPHGRKKNFVYVFTL 180
Qy 181 GQYFQKGRCSVRVSNVTANTVLGPQMEVTVYRSHGRAYPIAOVKDYVVTDOIPVF 240
Db 181 GQYFQKGRCSVRVSNVTANTVLGPQMEVTVYRSHGRAYPIAOVKDYVVTDOIPVF 240
Qy 241 TMFQKNDNSDETFLKPLIMFDVLIHDPSPHFLNYSINYNKWSFGDNTGLFVSNHTVN 300
Db 241 TMFQKNDNSDETFLKPLIMFDVLIHDPSPHFLNYSINYNKWSFGDNTGLFVSNHTVN 300
Qy 301 HTYVLNGTFTSLNLTVKAAAGPGCPPPPPPPR-----SKFTPSLGPAGDNP 346

```

```

Db 301 HTYVLNGTFTSLNLTVKAAAGPGCPSPSPSSSTSPSPASSPSPSTLSPSLMPTGYKS 360
Qy 347 LELSRIPDENQINRYGHFOATITIVGILRVNIITQMDVLPVWPWPRESSLIDFVVTQGG 406
Db 361 MELSDISNENCRINRYGVFRATITIVGILRVNIITQMDVLPVWPWPRESSLIDFVVTQGG 420
Qy 407 SIPTVECTIISDPTCQIAQNRVCSFVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 466
Db 421 ATPTACTIISDPTCQIAQNRVCSFVAVDELCLLSVRRAFNGSGTYCNFTLGGDASLAL 480
Qy 467 TSTLSVPRDRPASPLRMANSALISVGCLAFVTVISLLVYKHKYKYNPIENSPGNVRS 526
Db 481 TSALISIPGKDLGSLRTVNGVLISIGCLAFVTVISLLVYKHKYKYNPIENSPGNVRS 540
Qy 527 KGLSVFLNRAKAVFPFGNQEKDPLLNQ 554
Db 541 KGLSVFLSHAKAFPSRGDREKDPQLQDK 568

RESULT 5
Q8BVA0 MOUSE PRELIMINARY; PRT; 536 AA.
AC Q8BVA0 MOUSE PRELIMINARY; PRT; 536 AA.
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Mus musculus adult male urinary bladder cDNA, RIKEN full-length
DE enriched library, clone:9530038P20 product:glycoprotein
DE (transmembrane) nmb, full insert sequence.
GN Name=Gpmb;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleisemann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guetlich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RN Nature 409:685-690(2001).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Urinary bladder;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

```





FT CARBOHYD 92 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 133 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 145 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 149 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 192 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 199 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 248 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 274 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 307 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 311 N-linked (GlcNAc...) (Potential)  
FT CARBOHYD 459 N-linked (GlcNAc...) (Potential)  
SQ SEQUENCE 559 AA; 61714 MW; 8182793BD64A47DD CRC64;

Query Match 52.5%; Score 1588.5; DB 1; Length 559;  
Best Local Similarity 53.5%; Pred. No. 2.7e-111;  
Matches 297; Conservative 86; Mismatches 157; Indels 15; Gaps 5;

QY 8 LGFLLLAARLPDAAKRFHDLGNRPSAYMRHNLGWSNDNDWNEKLYPVWKRGM 67  
DB 8 LALLLPAAVLCAAAMRQDVLNQR-TAPVTHHKKIQGWSDDQKWEKLYPFWEDNDP 66  
QY 68 RWKSWKGRVQAVLTSDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLS 127  
DB 67 RKWCWKGGKVTTLVTDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLS 126  
QY 128 ADPYVNTANSESDGNGTCQSHNVFPDGPFPHPGWRNFIYVFTLQGYFQKL 187  
DB 127 QDQYVNTWETIDNCGWENCTSNHSHNVFPDGPFPHPGWRNFIYVFTLQGYFQKL 186  
QY 188 GRCSVRVSVNTANVTGLPQLEMTVYVRHGRAYVIAQVKDQVYVTDQIPVVTMFOKND 247  
DB 187 GRSSANFSVNTANITLGHKIMAVSIVRRGHSTVYPIARASTYVVTDKIPILVNSQKHD 246  
QY 248 RNSDTEFLKOLPIMFDVLIHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNTYVLNG 307  
DB 247 RNISDIIFKDSPIITFDVKIHDPSYLLNDSALSYKNFNGDGLFVSGATTSHTFSLOG 306  
QY 308 TFSMLTVKAAGPCPPPPPPRPSKTPSL-----GPAGNPLELSRIPDENCOI 359  
DB 307 NFTLNLTVQAIIPVPC--KPVTPPSLPTPAVTTDASSNSDPSAPNMAEDN-PDGGCHI 363  
QY 360 NRYGHFQATITVEGILEVNIQMTDVLMPVWPPESSILDFVTCQSGIPTVECTIISDP 419  
DB 364 YRYGYTTAGITVEGILEVNIQMTSIQMTSQAENPLVDFVVTQCGSFPDVCVANSDDP 423  
QY 420 TCEITQNTVCSVDVDEMLLTVRRTFNGSGTYCVNLTLGDDTSLALSTLISVPDRDPA 479  
DB 424 TCQVSGMVCVPVVTDECVLTIRAFDEPGTYCINILTGDDTSQALASALISV---NGG 480  
QY 480 SPLRMANGLISVGCLAIFVTVISLLVYKHKYKYNPIENSPGNVRSKGLSVFLNRAKAV 539  
DB 481 SSSGTTKGVIPLGLLAVFGAIGAFVLYKRYKQYKPIERSAGQAEQGLSAYVNFKAF 540  
QY 540 FPGNQEKDPLKQK 554  
DB 541 FPKSTERNPLKSK 555

RESULT 7  
Q4V805\_XENLA PRELIMINARY; PRT; 595 AA.  
AC Q4V805;  
DT 13-SEP-2005 (Tremblrel. 31, Created)  
DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)  
DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)  
DE Hypothetical protein (Fragment).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus; Xenopus.  
OX NCBI\_TaxID=8355;  
RN (1)  
RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Spleen;  
RA MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RT Richardson P.;  
RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391 (2002).  
RN (2)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Spleen;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN (3)  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Spleen;  
RX Klein S., Gerhard D.S.;  
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC097624; AAH97624.1; -; mRNA.  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 595 AA; 65314 MW; 06883CF91309D29F CRC64;

Query Match 47.3%; Score 1431.5; DB 2; Length 595;  
Best Local Similarity 47.7%; Pred. No. 2.1e-99;  
Matches 277; Conservative 86; Mismatches 165; Indels 53; Gaps 6;

QY 21 AAKFHDVLGNRPSAYMRHNLGWSNDNDWNEKLYPVWKRGMKNSWKGRVQA 80  
DB 19 AVKFPDVLGROPVGNRSHRHLGWSNDEKRYPSWKAGDSWEKCRGKVA 78  
QY 81 VLTSDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLSADPYVNTWASE 140  
DB 79 RLTSDSPALVGSNTTFAVNLIPRCQKEDANGNIYVKNCRNEAGLSADPYVNTWASE 136  
QY 141 DSDGNGTCQSHNVFPDGPFPHPGWRNFIYVFTLQGYFQKLGRCSVRVSVNTAN 200  
DB 137 -NDFCNEGCSFNKSPDGPFPHPGWRNFIYVFTLQGYFQKLGRCSVRVSVNTAN 195  
QY 201 VTLPQLEMTVYVRHGRAYVIAQVKDQVYVTDQIPVVTMFOKNDNSDDETLKDL 260  
DB 196 ITAQVQVSVYVRHGRAYVIAQVKDQVYVTDQIPVVTMFOKNDNSDDETLKDL 255  
QY 261 IMPDVLHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNTYVLNGTFSNLTVKAAAP 320  
DB 256 INFDIRIHDPHFNLSTYNTYKWSFGDNTGLFVSTNHTVNTYVLNGTFSNLTVKAAAP 315  
QY 321 GPCPP-----PPPPRPSKTP-----SLGAGDNPLELSRIP----- 353  
DB 316 SPCNPLATTPVTTFVTPVTPVTAQMPPTTTSPTAQTGNTDSEPALFTTEVPTEITTT 375  
QY 354 -----DENCQINRYGHFQATITVEGILEVNIQMTDVLMPVWPPE 394  
DB 376 TARTTTTARTTTTAAAGCFIYRYGYSTNTITVVDGIVEINIEMTNVQVPASQAE 435



```
QY 395 SSLIDFVTCGSIPTVEVTHIISDPTCEITONTVCSPVDVDEMCLLTVRTFNGSGTYCV 454
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
436 NALIDFVTCGSLPKDVCTLIISDTCMIPQNVCEEPATEQCSTLRRFAFAQPGTYCV 495
QY 455 NLITGDDTSALNSTLSVDPDRPASPLRMANSALISVGCLAIFVTVISILLVVKCKEYN 514
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
496 NITLSDASALASTLSV--SGSAKQKTVAAVPLPGFVIVAAGVAILLKKEYR 553
QY 515 PIENSPGNVRSKGLSVFLNRKAVFPFQNGEQDPLLNQK 555
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
554 PIDNAATS--DNRGLSVYFQVSKVSLFPNGNEQDPLLKTKQ 592

RESULT 8
Q96F58 HUMAN
ID Q96F58 HUMAN PRELIMINARY; PRT; 206 AA.
AC Q96F58
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Hypothetical protein GNMB (Glycoprotein (Transmembrane) nmb).
GN Name=GNMB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klatsner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dratchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,
RA Phelan M., Farmer A.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=99063792; PubMed=9847074;
RA Wilson R.;
RT "Toward a complete human genome sequence."
RL Genome Res. 8:1097-1108(1998).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Kalicki J., Kramer J.;
RT "The sequence of Homo sapiens BAC clone CTA-271G13."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
```

```

[6]
RN NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [7]
RP NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011595; AAH11595.1; -; mRNA.
DR EMBL; BT007074; AAP35737.1; -; mRNA.
DR EMBL; AC005082; AAP22337.1; -; Genomic DNA.
GO; GO:0016021; C:integral to membrane; IEA.
KW Hypothetical protein; Transmembrane.
SQ SEQUENCE 206 AA; 23945 MW; 30C8DE6928D73FBD CRC64;

Query Match 34.0%; Score 1029; DB 2; Length 206;
Best Local Similarity 98.4%; Pred. No. 1.4e-69;
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLNGWSSDENDWNEKLYP 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLNGWSSDENDWNEKLYP 60

QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIVYEKNC 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIVYEKNC 120

QY 121 RNEAGLSADPYVYNWTAWSGSDGNGTQSHNVFPDGPFPHPHGWRRWNIYVPHTL 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 RNEAGLSADPYVYNWTAWSGSDGNGTQSHNVFPDGPFPHPHGWRRWNIYVPHTL 180

QY 181 GOYFQ 185
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 GWLLQ 185

RESULT 9
Q81XJ5 HUMAN
ID Q81XJ5 HUMAN PRELIMINARY; PRT; 206 AA.
AC Q81XJ5
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Glycoprotein nmb-like protein.
GN Name=GNMB;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lernerz V.;
RT "Identification and Characterization of T cell defined tumor antigens
RL in the melanoma model MZ7."
RL Thesis (2002), Department of Biological Sciences, Johannes
RL Gutenberg-University of Mainz, Mainz, Germany.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Lernerz V., Facho M., Gentilini C., Lifke A., Woelfel C., Woelfel T.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ505015; CAD43718.1; -; mRNA.
DR Ensembl; ENSG00000136235; Homo sapiens.
SQ SEQUENCE 206 AA; 24003 MW; 30C8EE5A28D73FBD CRC64;

Query Match 33.8%; Score 1022; DB 2; Length 206;
Best Local Similarity 97.8%; Pred. No. 4.7e-69;
Matches 181; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLNGWSSDENDWNEKLYP 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMREHNLNGWSSDENDWNEKLYP 60
```

```

QY 61 VVKRDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC 120
Db 61 VVKRDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC 120
QY 121 RNEAGLSADPYVYNNWTANSESDGNGTQSHNVFPDQKPPHPGRWRNFIYVHTL 180
Db 121 RNEAGLSADPYVYNNWTANSESDGNGTQSHNVFPDQKPPHPGRWRNFIYVHTL 180
QY 181 GOYFQ 185
Db 181 DWLLQ 185

RESULT 10
Q4KX9 TETNG
ID Q4KX9 TETNG PRELIMINARY; PRT; 538 AA.
AC Q4KX9
DT 13-SEP-2005 (TremBLrel. 31, Created)
DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)
DE Chromosome 21 SCAP15029, whole genome shotgun sequence.
GN ORFNames=GSTENG00032949001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]_
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bounau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA Bionnet C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McSwan P., Bosak S.,
RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAZ01015029; CAG11103.1; -; Genomic DNA.
SQ SEQUENCE 538 AA; 60051 MW; EAA1B2548813D9E CRC64;

Query Match 27.9%; Score 843; DB 2; Length 538;
Best Local Similarity 36.0%; Pred. No. 5.5e-55;
Matches 191; Conservative 84; Mismatches 214; Indels 42; Gaps 11;

QY 46 GWSNDNWEKLYPV--WKQDM-RWNSWKGVRVQAVLTSDSPALVGSNITFAVNLIF 102
Db 46 GWSNDNWEKLYPV--WKQDM-RWNSWKGVRVQAVLTSDSPALVGSNITFAVNLIF 102
QY 20 GWDQTNPDWDDLYPLNLKTNLTRHK-----GKPKVHLTSDSPALVGSNITFAVNLIF 74
Db 20 GWDQTNPDWDDLYPLNLKTNLTRHK-----GKPKVHLTSDSPALVGSNITFAVNLIF 74
QY 103 PRCKEDANGNIYKNCNEAGLSADPYVYNNWTANSESDGNGTQSHNVFPDQKPP 162
Db 103 PRCKEDANGNIYKNCNEAGLSADPYVYNNWTANSESDGNGTQSHNVFPDQKPP 162
QY 75 PPCKEDAGGDDWDEHCEHEDANGQIRSGVYNNWTANSESDGNGTQSHNVFPDQKPP 134
Db 75 PPCKEDAGGDDWDEHCEHEDANGQIRSGVYNNWTANSESDGNGTQSHNVFPDQKPP 134
QY 163 PHHPGRWRNFIYVHTLGOYFQKLGRCGRSVRVSVANTVTLGPOLMEVTVYR-RHGRAYV 221
Db 163 PHHPGRWRNFIYVHTLGOYFQKLGRCGRSVRVSVANTVTLGPOLMEVTVYR-RHGRAYV 221
QY 135 PQSNDWRHKSYYVWHTWQYETCDGSSSVYVNTHTPIGAEIMEVLVYKRRKYS 194
Db 135 PQSNDWRHKSYYVWHTWQYETCDGSSSVYVNTHTPIGAEIMEVLVYKRRKYS 194
QY 222 PIAQVKDYVVTDDQIPVFTVYVQKNDNRSSDFTFLKDLPIMPDVLIDHPSHFL-NYSTIN 280
Db 222 PIAQVKDYVVTDDQIPVFTVYVQKNDNRSSDFTFLKDLPIMPDVLIDHPSHFL-NYSTIN 280

```

```

Db 195 PLTVNTVYVYTDKIPAVSISQKAAVNSQSENVPFRGKDVVKVQLHDPDGVLTAAALD 254
QY 281 YKWSFGDNTGLFVSTNHTVNTYVNLNGTFSNLNTVKAAPGCPPPPP-----PPRPSK 334
Db 255 YIWDPRDGNQVLTNRVNT--THYSRLGTMSVKLVVEAAFPABCPPTSAPATERLPPTVS 313
QY 335 PTPSLGPAGDNPLESRIIPDENQCINRYGHQATITIVIGILEVNIQWTDVLMVPWPPE 394
Db 314 ETPESGPTTLDWLRVRLRSRQCPRYVHGTFVGNITITIEPPQLKSA-----PH 362
QY 395 SSLID-----FVWTCOGSIPTEVCTIISDPTCEITONTVCSVDVDEMCLLTV 442
Db 363 SRIVDVASRVTKTDVSVFLVKLGNIPVSACTIVSDPCTEVRAITCDDVPPLSECEVRL 422
QY 443 RRTFNGSGTYCVNLTLGDDTSLATSTLISV-PDRD-PASPLRMANSALISVGLCIAFVT 500
Db 423 QRSFAEPCTYCVNITLEDGSGVALASTVTVINKSQDAPETSSRAAALSSAAVLVVFA 482
QY 501 VISLLVYKHKYENPIENSPGNVRSKLSVFLNRAKAVFPFGNOEKDPLL 551
Db 483 FAAYLVCRRYKGYRVPORPALGSCICRYGVGGVGHVHLREKLPFPRESHILL 533

RESULT 11
Q4KLV9 XENLA
ID Q4KLV9 XENLA PRELIMINARY; PRT; 719 AA.
AC Q4KLV9
DT 13-SEP-2005 (TremBLrel. 31, Created)
DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TremBLrel. 31, Last annotation update)
DE LOC445868 protein (Fragment).
GN Name=LOC445868;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RC MEDLINE=2341132; PubMed=12454917; DOI=10.1002/dvdy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.,
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative."
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Statchenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.C., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;

```

```
RA Klein S., Gerhard D.S.;
RL Submitted (JUL-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC098976; AAH98976.1; -, mRNA.
FT NON TER
SQ SEQUENCE 719 AA; 76930 MW; 54A4D018A028C9E3 CRC64;

Query Match      24.4%; Score 738.5; DB 2; Length 719;
Best Local Similarity 26.4%; Pred. No. 6.4e-47;
Matches 189; Conservative 97; Mismatches 206; Indels 225; Gaps 19;

Qy 43 QLNWSSDE-----NDWEKLYPVKRGDMRWKNSKGGRVQAVLTSDS 86
Db 21 QTONWQNRVQQTQQVAGNRQSPFKWNSMYPIMRGSAQKNCWKGGQVTPDLVYDA 80
Qy 87 PALVGSNTITFAVNLIFPRCKEDANGNIIVVEKKNCRNAG--LSADPVYVNWTAWSDDG 144
Db 81 PTLTGAKATFSIRLNFNNQVLPDGGVWQGN-RTONGTWIPSEEIYP-----DESTEG 135
Qy 145 ENGTGQSHHNVPDGKPFPHHGWRRWNIYVFTLQGYFQKLGRCSSVRVSVNTANTVLG 204
Db 136 SECT-----FPDGRFPFRGVEKKHSKFVYVMTGWKYVQVDPGSSNLTVTETDGIPLG 188
Qy 205 POLMEVTVYRRHGR-AVYPIAQKDVVVTVDQIPVFTFMQKNDNRNSDETFLKDLIMF 263
Db 189 SYTMQVVVYHYRGQRKQIPIGSISQFTITDQIPVSVSISQLLDQEDQRFIQNRVSVF 248
Qy 264 DVLLHDPHSFLNYSTINVKNSFGDNTGLFVSTNHTVNHVTYVNGTFSNLTVKAAAP-GP 322
Db 249 AVAIDHPSHYLQAADISFSWDFGQSGTILTRNDVTHTYVSPGVFRPKVQLQAIPAP 308
Qy 323 C-----PPPPPPPRPSKPTPSLG-----PAG----- 343
Db 309 CGSTAPVATAEPVTTVAPAQPTTAAASTSGGVTDAPPCTTAALPGNITEPQTINGIV 368
Qy 344 -----DN-----PLELSRIPDE----- 355
Db 369 VTIPSDEQNLLAEATLSSONEAATLPEAIDEAGVTVAGSETVPEEEAVPNOEQAVTV 428
Qy 356 ----- 355
Db 429 ABAVPNOEQAVTVPNQOAVTVAEAVPSQBAEAVTSEATTEATAELARGSPSVPSQINE 488
Qy 356 -----NCQINRYGHFQATITTVEGILEVNI 381
Db 489 LPGTTAEVLTVPAEAAEEVVVIAKQAPEDPLVGLLYRGTFATDLDIVQGLSAQIV 548
Qy 382 QMTDVLMPVPWPESSLIDFVTCOGSIPTEVCTTIISDPTCEITQNTVCSPVDVDEMCLLT 441
Db 549 QVAPV-AAVDGAENA-VDLATTCOGSVPSQVCTTIISNPDCPTIPOETVCNTVQPSTDCQLV 606
Qy 442 VRRTFNGSGTVCVNLTGDDTSLATLTSLISVPDRDPASPLRMANSALISVG-CLAIPTV 500
Db 607 LRQVFNTEGLVCVNVSLTSDVSLMASTQVSVSGAGSS----FSGGIVIVGVLLAVFAV 662
Qy 501 VISLLVYKCKHEYNPNIEN--SPGNVWRSK-GLSVFLNRKAVFPFGNQEKDPLLNQ 554
Db 663 VIVATYTRQKSYALTATAPSPSNWFPDRLSLRLFFQNALGLSRG--ENSPLNGR 717

RESULT 12
Q6DDN6 XENLA
ID Q6DDN6 XENLA PRELIMINARY; PRT; 721 AA.
AC Q6DDN6
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE LOC445868 protein (Fragment).
GN Name=LOC445868;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
```

```
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC MEDLINE=22341112; PubMed=12454917; DOI=10.1002/dvdy.10174;
RX Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RL Dev. Dyn. 225:384-391(2002).
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shennen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN NUCLEOTIDE SEQUENCE.
RP TISSUE=Ovary;
RC Klein S., Gerhard D.S.;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC077508; AAH77508.1; -, mRNA.
DR InterPro; IPR006001; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
FT NON TER
SQ SEQUENCE 721 AA; 77121 MW; 9567A3BADD205F78 CRC64;

Query Match      24.3%; Score 736.5; DB 2; Length 721;
Best Local Similarity 27.0%; Pred. No. 9.1e-47;
Matches 186; Conservative 96; Mismatches 199; Indels 209; Gaps 18;

Qy 54 WNEKLYPVKRGDMRWKNSKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGN 113
Db 50 WNSRMYPIMRGSAQKNCWKGGQVTPDLVNDAPTILGAKATFSIRLNFNNQVLPDQG 109
Qy 114 IVYEKKNCRNAG--LSADPVYVNWTAWSDDGNGTQSHHNVPDGKPFPHHGWRRW 171
Db 110 VVWQGN-RTONGTWIPSEEIYP-----DESTGSECT-----FPDGRFPFRGVEKKHS 157
Qy 172 NFIVFHTLQGYFQKLGRCSSVRVSVNTANTVLGQLMEVTVYRRHGR-AVYPIAQKDVY 230
Db 158 KFYVVMQTWKYVQVVDGSPSSNLTVTETDGIPLGSIYTMQVVVYHYRGQRKQIPIGSISQF 217
Qy 231 VVTDQIPVFTFMQKNDNRNSDETFLKDLIPMDVLHDPHSFLNYSTINVKNSFGDNTG 290
Db 218 TITDQIPVSVSISQLLDQEDQRFIQNRVAVFAVIAHDPSHYLQAADISFSWDFGDSG 277
Qy 291 LFSVTNHTVNHVTYVNGTFSNLTVKAAAP-GPC-----PPPPPPPRPSKPTPSLG 340
Db 278 TLITRNDVTHTYVSPGVFRPKVQLQAIPAPCGSTAPVATAEPVTTVAPAQPTTAA 337
Qy 341 -----PAG-----DN----- 345
Db 338 STSGGVTDAPPCTTAALPGNITEPQTINGIVVVTIPSDEQNLLAEATLSSDNEAATLP 397
```

QY 346 -----PLESRIDP----- 355  
 Db 398 EAIEDAGVTAGBETVPEEAEVNPQEQAVTVAEAVNPQEQAVTVAEAVP 457  
 QY 356 ----- 355  
 Db 458 SQEAEVTSAEETAEIAELAGSPVPSQINELPGTTAEVLTVPATERAEVVIARQ 517  
 QY 356 -----NQINRYGHFOATITVIGILEVNIQMTDVLMPVPPESSLIDFVTCQSI 408  
 Db 518 APEDPLVGLLYRGTFATDLIVOGIESAQIVQVAPV-AAVDGAENA-VDLTITCQGSV 575  
 QY 409 PTEVCTIISDPCEITONTVCSPVDENCLLTVRTFNGSGTYCNLTGLDTSIALTS 468  
 Db 576 PSQVCTIISNPCTIPQETVNTVPSTDCQLVLRQVFNETGLYCNVSLTDSVLAMAS 635  
 QY 469 TLISVDPDRPASFLMANSALISVG-CLAIFVTVLSLVYKHKYENPIEN--SPGNVVR 525  
 Db 636 TQVSVSGAGSS-----FSNGGIVIVGVLAVFAVIVAYTYRQKSYTALTAPSNSWFP 691  
 QY 526 SK-GLSVFLNRAKAVFPFGNQEKDPLLNQ 554  
 Db 692 DRLSLRPLFONALGLSRSG--ENSPLNGR 719

RESULT 13  
 Q6DIR2\_XENTR  
 ID Q6DIR2\_XENTR PRELIMINARY; PRT; 746 AA.  
 AC Q6DIR2  
 DT 25-OCT-2004 (TRENBLrel. 28, Created)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
 DE SILV protein (Fragment).  
 GN Name=SILV;  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Silurana.  
 OX NCBI\_TaxID=8364;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo.  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide P.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Falley J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RP Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC TISSUE=Embryo;  
 RA Klein S., Gerhard D.S.;  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC075473; AAH75473.1; -; mRNA.  
 DR InterPro: IPR000601; PKD.  
 DR Pfam: PF00801; PKD; 1.  
 DR SMART: SM00089; PKD; 1.

DR PROSITE, PSS0093; PKD; 1.  
 FT NON\_TER 1  
 SQ SEQUENCE 746 AA; 79230 MW; D2D3AB14A134D349 CRC64;  
 Query Match 24.1%; Score 729.5; DB 2; Length 746;  
 Best Local Similarity 26.6%; Pred. No. 3.2e-46;  
 Matches 197; Conservative 92; Mismatches 204; Indels 247; Gaps 19;  
 QY 28 VLGNRPSPAYMREHNQNLNGSSDENWNEKLYPVWKRGMWNSWKGSRVQAVLTSSDP 87  
 Db 39 VAGNRQPPF-----KSMNSRMYPIMRGTEAKKCKWCKGQVTFNLVNDAP 83  
 QY 88 ALVGSNITFAVNLIPPCQKEDANGNIYVKNCRNEAGL---SADPVYNNWTAMSEDSG 144  
 Db 84 TLTGAKATFQLNFPKQNTVLPQGVVWGQ-RTDNGTWVPSGDP-VTP-----DESTG 137  
 QY 145 ENGTCQSHNVFPDCKPFPHPGWRMNFIVVFTLQYFQKLGRCSCVRVSVNTANVTIG 204  
 Db 138 SECT-----FPDGRFPFGSEKSKFVYVWQTLGKYWQVVDGPPSSNLTLETGIPLG 190  
 QY 205 POLMEVTVYRHRGR-AVYPIAQVKDVVYVTDQIPVFTVMPQKDRNSDETFLKOLPIMF 263  
 Db 191 SYTMQVVVHYRGROKFIPIGSASQFTITDQIPVSVSISQLDLQEDQRFIONRAVSF 250  
 QY 264 DVLHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVNTHTYVNLGTFSLNLTVYKAAAP-GP 322  
 Db 251 AVSIHDPSHYLAQADISFSDFGQSGTLITRNTDVTHTYVSPGVFRPKVVLQAAIPITP 310  
 QY 323 C-----PPP-----PPPP----- 331  
 Db 311 CGSTAAAGTAEPVSTVTPPQTTAVSTSGQVTDATGTPAVLPGNVTEPQGTINIV 370  
 QY 332 ---PSKPT----- 336  
 Db 371 VTLPEETQNNLAABATSSVSASALPDNEAATLPSGTBEAGVTDAABETVPEEVAEN 430  
 QY 337 -----PSLG----- 340  
 Db 431 QEQAABAVPSQAGVTVABAVPSQEAABAVPSQEAABAVPSQEAABAVPSQEAABAVT 490  
 QY 341 -----PAGDNPLELSRI-----PDEN 356  
 Db 401 SEAGNTEAGAEAEETPSVASEINELPGSTAEVTVTPATEGEEVVVIAKRAQAPDLVG 550  
 QY 357 COINRYGHFOATITVIGILEVNIQMTDVLMPVPPESSLIDFVTCQSIPTVCTII 416  
 Db 551 CLLRYGTGTFATDLIVOGIESAQIVQVAPV-AAVDGTENA-VDLTVCQGSVPVQVCTII 608  
 QY 417 SDPTCEITONTVCSPVDENCLLTVRTFNGSGTYCNLTGLDTSIALTSILSVPR 476  
 Db 609 SGPDCTPQETICNTVPSTDCQLVLRQVFNDETGLYCNVSLTDAVSLAMASTQVSVSGA 668  
 QY 477 DPASPLRMANSALISVG-CLAIFVTVLSLVYKHKYENPIENS-PGNVVRSGLSVFLN 534  
 Db 669 GSS-----FANGGIVIVGVLAVFAVIVAYYRQKSYTALTAPSNSWFP 724  
 QY 535 RAKAVFPFGNQEKDPLLNQ 554  
 Db 725 FQNALGLSRSGNSPLNGR 744

## RESULT 14

Q52K24\_XENLA  
 ID Q52K24\_XENLA PRELIMINARY; PRT; 808 AA.  
 AC Q52K24  
 DT 13-SEP-2005 (TRENBLrel. 31, Created)  
 DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)  
 DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus; Xenopus.





```

Qy 323 CPPPPP-----PPPS-----K 334
Db 339 MPTTQPSGTTVVQMPTEIATTSQMLTSAVIDTTLAENVSTTEGTGTTPTRPSTGTTVAQ 398
Qy 335 PTPSLGPAGDNPL-----ELSRIPDE-----NCQINRYGHFQATITIVE 373
Db 399 ATTEGPDGDA-SPLLFTOSSTGSI SPLLDGDTDTIMLVKRQVPLDCVLYRYGSFSLALDIVQ 457
Qy 374 GILEVNIITQMTDVLMPVWPPESSLIDFVVTCCGSIPTTEVCTIISDPTCEITQNTVCSPVD 433
Db 458 G-----IESAEIILQAVPFSEGDAPFELTVSCQGLPKXACMDISSPGCQPPAQRLCOSVP 511
Qy 434 VDEMCLLTVRRTF-NGSGTYCVNLTGLGDDTSLALTSTLISVPDRDPASPLRMANSALISV 492
Db 512 PSPDCQLVHLQVLKGGSGTYCLNVSLADANSLAVASTQLVVPQDGG-----LCQAPLLV 566
Qy 493 GCLAIPTVTVISLLVYKXH--KEYNPIENSPGN-----VVRSGLSVFLNRAKAVFF 541
Db 567 GILLVAVVLA SLIHRRLKKGVSQMPHGSTHRLRLPVFRARGIG----- 615
Qy 542 PGNOEKDPLAKNOE 555
Db 616 ----ENSPLLSGQQ 625

```

Search completed: December 24, 2005, 04:35:47  
Job time : 243 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:28:24 ; Search time 49 Seconds  
(without alignments)  
944.865 Million cell updates/sec

Title: US-09-853-880A-17  
Perfect score: 3026  
Sequence: 1 MECLYFGLLLAARLPD.....PFGNQKDPLLNQKFGKVS 560

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5 COMB pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6 COMB pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H COMB pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS COMB pep.\*  
5: /cgn2\_6/ptodata/1/iaa/R8 COMB pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	1 US-08-594-031-90	Sequence 90, Appl
2	3026	100.0	560	2 US-09-643-597-225	Sequence 225, App
3	3026	100.0	560	2 US-09-480-884A-225	Sequence 225, App
4	3026	100.0	560	2 US-09-542-615A-225	Sequence 225, App
5	3026	100.0	560	2 US-09-606-421B-225	Sequence 225, App
6	3026	100.0	560	2 US-09-476-496A-225	Sequence 225, App
7	3026	100.0	560	2 US-09-630-940B-225	Sequence 225, App
8	3026	100.0	560	2 US-09-943-075A-6	Sequence 6, Appl
9	3026	100.0	560	2 US-10-039-272A-2	Sequence 2, Appl
10	3026	100.0	560	2 US-10-007-700-225	Sequence 225, App
11	3026	100.0	560	5 US-09-985-799-90	Sequence 90, Appl
12	3026	100.0	560	5 US-09-977-371-90	Sequence 36, Appl
13	2197	72.6	574	2 US-09-383-586-36	Sequence 36, Appl
14	2197	72.6	574	2 US-09-823-038A-36	Sequence 36, Appl
15	2197	72.6	574	2 US-09-943-075A-5	Sequence 5, Appl
16	2179	72.0	572	2 US-09-197-970B-5	Sequence 2, Appl
17	2163	71.5	572	2 US-09-943-075A-2	Sequence 100, App
18	969	32.0	376	1 US-08-594-031-100	Sequence 102, App
19	969	32.0	376	1 US-08-594-031-102	Sequence 102, App
20	969	32.0	376	5 US-09-985-799-100	Sequence 100, App
21	969	32.0	376	5 US-09-985-799-102	Sequence 102, App
22	969	32.0	376	5 US-09-977-371-100	Sequence 102, App
23	969	32.0	376	5 US-09-977-371-102	Sequence 102, App
24	622	20.6	661	1 US-08-417-174-121	Sequence 121, App
25	622	20.6	661	2 US-09-267-439-121	Sequence 121, App
26	622	20.6	661	2 US-08-388-852B-2	Sequence 2, Appl
27	622	20.6	661	2 US-09-073-138-121	Sequence 121, App

Query Match 100.0%; Score 3026; DB 1; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;

28	622	20.6	661	2 US-09-862-260A-2	Sequence 2, Appl
29	621	20.5	661	1 US-08-417-174-27	Sequence 27, Appl
30	621	20.5	661	1 US-08-231-565A-27	Sequence 27, Appl
31	621	20.5	661	1 US-09-007-961-27	Sequence 27, Appl
32	621	20.5	661	2 US-09-267-439-27	Sequence 27, Appl
33	621	20.5	661	2 US-09-073-138-27	Sequence 27, Appl
34	603.5	19.9	668	1 US-07-891-942G-6	Sequence 6, Appl
35	354.5	11.7	460	2 US-09-949-016-8029	Sequence 8029, Ap
36	238.5	7.9	190	2 US-08-388-852B-35	Sequence 35, Appl
37	236	7.8	202	2 US-08-388-852B-38	Sequence 38, Appl
38	233.5	7.7	192	2 US-08-388-852B-37	Sequence 37, Appl
39	233	7.7	197	2 US-08-388-852B-36	Sequence 36, Appl
40	181.5	6.0	423	2 US-09-991-181-241	Sequence 241, App
41	181.5	6.0	423	2 US-09-990-444-241	Sequence 241, App
42	181.5	6.0	423	2 US-09-997-333-241	Sequence 241, App
43	181.5	6.0	423	2 US-09-992-598-241	Sequence 241, App
44	136.5	4.5	71	1 US-08-594-031-98	Sequence 98, Appl
45	136.5	4.5	71	5 US-09-985-799-98	Sequence 98, Appl

ALIGNMENTS

RESULT 1  
US-08-594-031-90  
; Sequence 90, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; FILING DATE: 30-JAN-1996  
; APPLICATION NUMBER: US/08/594,031  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 90:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
US-08-594-031-90

Query Match 100.0%; Score 3026; DB 1; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;

Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYFGLFLLAARLPDAAKGFHDVNGRPSAYMRHNLQNGSSDENDWNEKLYP 60  
Db 1 MECLYFGLFLLAARLPDAAKGFHDVNGRPSAYMRHNLQNGSSDENDWNEKLYP 60  
Qy 61 VKRGDMWRKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVEKNC 120  
Db 61 VKRGDMWRKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVEKNC 120  
Qy 121 RNEAGLSADPYVYNWTAWSESDGNGTQSHHNVFPDGKPPHHPGWRWNFIYVFHTL 180  
Db 121 RNEAGLSADPYVYNWTAWSESDGNGTQSHHNVFPDGKPPHHPGWRWNFIYVFHTL 180  
Qy 181 GOYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240  
Db 181 GOYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240  
Qy 241 TMFQKNDRNSDETFLKDLPIMFVDLIHDPESHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
Db 241 TMFQKNDRNSDETFLKDLPIMFVDLIHDPESHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCQIN 360  
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCQIN 360  
Qy 361 RYGHFOATITVEGILEVNIQMTDVLMPVPWPBSSLIDFVTCQGSIPTEVCIIISDPT 420  
Db 361 RYGHFOATITVEGILEVNIQMTDVLMPVPWPBSSLIDFVTCQGSIPTEVCIIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSLISVDRDPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSLISVDRDPAS 480  
Qy 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVRSKGLSVFLNRAKAVF 540  
Qy 541 PPGNQEKDPLLKNQEFKGS 560  
Db 541 PPGNQEKDPLLKNQEFKGS 560

RESULT 2  
US-09-643-597-225  
; Sequence 225, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-643-597-225

Query Match 100.0%; Score 3026; DB 2; Length 560;

Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYFGLFLLAARLPDAAKGFHDVNGRPSAYMRHNLQNGSSDENDWNEKLYP 60  
Db 1 MECLYFGLFLLAARLPDAAKGFHDVNGRPSAYMRHNLQNGSSDENDWNEKLYP 60  
Qy 61 VKRGDMWRKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVEKNC 120  
Db 61 VKRGDMWRKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVEKNC 120  
Qy 121 RNEAGLSADPYVYNWTAWSESDGNGTQSHHNVFPDGKPPHHPGWRWNFIYVFHTL 180  
Db 121 RNEAGLSADPYVYNWTAWSESDGNGTQSHHNVFPDGKPPHHPGWRWNFIYVFHTL 180  
Qy 181 GOYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240  
Db 181 GOYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQVKDVVVVTDQIPVFV 240  
Qy 241 TMFQKNDRNSDETFLKDLPIMFVDLIHDPESHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
Db 241 TMFQKNDRNSDETFLKDLPIMFVDLIHDPESHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
Qy 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCQIN 360  
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGNPLELSRIPDENCQIN 360  
Qy 361 RYGHFOATITVEGILEVNIQMTDVLMPVPWPBSSLIDFVTCQGSIPTEVCIIISDPT 420  
Db 361 RYGHFOATITVEGILEVNIQMTDVLMPVPWPBSSLIDFVTCQGSIPTEVCIIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSLISVDRDPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSLISVDRDPAS 480  
Qy 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVRSKGLSVFLNRAKAVF 540  
Qy 541 PPGNQEKDPLLKNQEFKGS 560  
Db 541 PPGNQEKDPLLKNQEFKGS 560

RESULT 3  
US-09-480-884A-225  
; Sequence 225, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-480-884A-225

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYFGLFLLAARLPDAAKGFHDVNGRPSAYMRHNLQNGSSDENDWNEKLYP 60  
|||||

Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPYVYNWTAMSESDGSHHNVFPDGKPPHHPGRRWNFIYVFHTL 180  
Db 121 RNEAGLSADPYVYNWTAMSESDGSHHNVFPDGKPPHHPGRRWNFIYVFHTL 180  
QY 181 GOYFQKLGRCVSRVSVNTANVTGLQPMFVLIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
Db 181 GOYFQKLGRCVSRVSVNTANVTGLQPMFVLIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
QY 241 TMFQKNDNRSSDETFKDLPIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
Db 241 TMFQKNDNRSSDETFKDLPIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
QY 301 HTYVLNGTFSNLTAKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360  
Db 301 HTYVLNGTFSNLTAKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360  
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVVTCCGSIPTVECTIISDPT 420  
Db 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVVTCCGSIPTVECTIISDPT 420  
QY 421 CSITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
Db 421 CSITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
QY 481 PLRMANSAISVGCIAIFVTYISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db 481 PLRMANSAISVGCIAIFVTYISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
QY 541 PFGNQEKDPLLNQKQEFKGS 560  
Db 541 PFGNQEKDPLLNQKQEFKGS 560

## RESULT 4

US-09-542-615A-225  
; Sequence 225, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-542-615A-225

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPYVYNWTAMSESDGSHHNVFPDGKPPHHPGRRWNFIYVFHTL 180  
Db 121 RNEAGLSADPYVYNWTAMSESDGSHHNVFPDGKPPHHPGRRWNFIYVFHTL 180  
QY 181 GOYFQKLGRCVSRVSVNTANVTGLQPMFVLIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
Db 181 GOYFQKLGRCVSRVSVNTANVTGLQPMFVLIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
QY 241 TMFQKNDNRSSDETFKDLPIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
Db 241 TMFQKNDNRSSDETFKDLPIHDPSPHFLNYSNTYKWSFGDNTGLFVSNHTVN 300  
QY 301 HTYVLNGTFSNLTAKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360  
Db 301 HTYVLNGTFSNLTAKAAAGPCPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360  
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVVTCCGSIPTVECTIISDPT 420  
Db 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPESSLIDFVVTCCGSIPTVECTIISDPT 420  
QY 421 CSITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
Db 421 CSITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPRDRPAS 480  
QY 481 PLRMANSAISVGCIAIFVTYISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db 481 PLRMANSAISVGCIAIFVTYISLLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
QY 541 PFGNQEKDPLLNQKQEFKGS 560  
Db 541 PFGNQEKDPLLNQKQEFKGS 560

## RESULT 5

US-09-606-421B-225  
; Sequence 225, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT APPLICATION NUMBER: US/09/606,421B  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-606-421B-225

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

QY 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHGRRWRFYVVFHTL 180  
DB 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHGRRWRFYVVFHTL 180  
QY 181 GOYFQKLGRCRSVRVSVNTANVTGLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQI1PVFV 240  
DB 181 GOYFQKLGRCRSVRVSVNTANVTGLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQI1PVFV 240  
QY 241 TMFQKNDNRNSDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300  
DB 241 TMFQKNDNRNSDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGNPLELSRIPDENCOIN 360  
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGNPLELSRIPDENCOIN 360  
QY 361 RYGHFOATITIVEGILEVNI1QMTDVLMPVPWPBESSLIDFVVTCCGSIPTIEVCTIIISDPT 420  
DB 361 RYGHFOATITIVEGILEVNI1QMTDVLMPVPWPBESSLIDFVVTCCGSIPTIEVCTIIISDPT 420  
QY 421 CEITQNTVCSFVVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480  
DB 421 CEITQNTVCSFVVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480  
QY 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540  
QY 541 FPGNQEKDPLLNKQEFKGV5 560  
DB 541 FPGNQEKDPLLNKQEFKGV5 560

## RESULT 6

US-09-476-496A-225  
; Sequence 225, Application US/09476496A  
; Patent No. 6706262

GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong

; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF

; TITLE OF INVENTION: LUNG CANCER  
; FILE REFERENCE: 210121.455C5

; CURRENT APPLICATION NUMBER: US/09/476,496A  
; CURRENT FILING DATE: 1999-12-30

; NUMBER OF SEQ ID NOS: 254  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 225  
; LENGTH: 560

; TYPE: PRT  
; ORGANISM: Homo sapien

US-09-476-496A-225

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRP5AYMRHNLQNGWSSDENDWNEKLYP 60  
DB 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRP5AYMRHNLQNGWSSDENDWNEKLYP 60  
QY 61 VWKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120  
DB 61 VWKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120  
QY 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHGRRWRFYVVFHTL 180  
DB 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHGRRWRFYVVFHTL 180  
QY 181 GOYFQKLGRCRSVRVSVNTANVTGLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQI1PVFV 240

DB 181 GOYFQKLGRCRSVRVSVNTANVTGLPQMEVTVYRRHGRAYVPIAQVKDVVVVTDQI1PVFV 240  
QY 241 TMFQKNDNRNSDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300  
DB 241 TMFQKNDNRNSDETFKDLPIHDPVLIHDPHFLNYSTINYNKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGNPLELSRIPDENCOIN 360  
DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPSPKPTPSLGPAGNPLELSRIPDENCOIN 360  
QY 361 RYGHFOATITIVEGILEVNI1QMTDVLMPVPWPBESSLIDFVVTCCGSIPTIEVCTIIISDPT 420  
DB 361 RYGHFOATITIVEGILEVNI1QMTDVLMPVPWPBESSLIDFVVTCCGSIPTIEVCTIIISDPT 420  
QY 421 CEITQNTVCSFVVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480  
DB 421 CEITQNTVCSFVVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDRDPAS 480  
QY 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGCIAIFVTVISLLVYKHKYENPIENSPGNVVRSGLSVFLNRAKAVF 540  
QY 541 FPGNQEKDPLLNKQEFKGV5 560  
DB 541 FPGNQEKDPLLNKQEFKGV5 560

## RESULT 7

US-09-630-940B-225  
; Sequence 225, Application US/09630940B  
; Patent No. 6737514

GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy

; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.

; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.455C10  
; CURRENT APPLICATION NUMBER: US/09/630,940B

; CURRENT FILING DATE: 2000-08-02  
; NUMBER OF SEQ ID NOS: 367

; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 225

; LENGTH: 560  
; TYPE: PRT

; ORGANISM: Homo sapien  
US-09-630-940B-225

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRP5AYMRHNLQNGWSSDENDWNEKLYP 60  
DB 1 MECLYYFLGFLLLAARLPDAAKRFHVLGNRP5AYMRHNLQNGWSSDENDWNEKLYP 60  
QY 61 VWKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120  
DB 61 VWKRGDMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYKNC 120  
QY 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHGRRWRFYVVFHTL 180  
DB 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHGRRWRFYVVFHTL 180

QY 181 GYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240  
DB 181 GYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240  
QY 241 TFMQKNDNSDETFLKDLPIFMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TFMQKNDNSDETFLKDLPIFMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLTAKAAAPGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
DB 301 HTYVLNGTFSNLTAKAAAPGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
QY 361 RYGHFOATITIVGEGILEVNIQMTDVLMPVPMPPESSLDIFVVTCCGSIPTVECTIISDPT 420  
DB 361 RYGHFOATITIVGEGILEVNIQMTDVLMPVPMPPESSLDIFVVTCCGSIPTVECTIISDPT 420  
QY 421 CEITQNTVCSVPDDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480  
DB 421 CEITQNTVCSVPDDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480  
QY 481 PLRMANSALISVGLCLAIPTVYISLLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGLCLAIPTVYISLLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAVF 540  
QY 541 PPGNQEKDPLLKNQBFKGV 560  
DB 541 PPGNQEKDPLLKNQBFKGV 560

## RESULT 8

US-09-943-075A-6  
; Sequence 6, Application US/09943075A  
; Patent No. 6812002  
; GENERAL INFORMATION:  
; APPLICANT: Popoff, Steven N.  
; APPLICANT: Safado, Payez P.  
; APPLICANT: Owen, Thomas A.  
; APPLICANT: Smock, Steven L.  
; TITLE OF INVENTION: Osteocalcin Protein and Nucleic Acids Encoding the Same,  
; FILE REFERENCE: 71369.262  
; CURRENT APPLICATION NUMBER: US/09/943,075A  
; PRIOR FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: US 60/229,006  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Human  
US-09-943-075A-6

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
DB 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
QY 61 VMKGRDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
DB 61 VMKGRDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPYNNWTAMSESDGNGTQSHHNPFGKPPHHPGWRWNFIYVFHTL 180  
DB 121 RNEAGLSADPYNNWTAMSESDGNGTQSHHNPFGKPPHHPGWRWNFIYVFHTL 180  
QY 181 GYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240  
DB 181 GYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240

QY 241 TFMQKNDNSDETFLKDLPIFMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TFMQKNDNSDETFLKDLPIFMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLTAKAAAPGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
DB 301 HTYVLNGTFSNLTAKAAAPGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
QY 361 RYGHFOATITIVGEGILEVNIQMTDVLMPVPMPPESSLDIFVVTCCGSIPTVECTIISDPT 420  
DB 361 RYGHFOATITIVGEGILEVNIQMTDVLMPVPMPPESSLDIFVVTCCGSIPTVECTIISDPT 420  
QY 421 CEITQNTVCSVPDDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480  
DB 421 CEITQNTVCSVPDDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480  
QY 481 PLRMANSALISVGLCLAIPTVYISLLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGLCLAIPTVYISLLVYKHKKEYNPPIENSPGNVRSKGLSVFLNRAKAVF 540  
QY 541 PPGNQEKDPLLKNQBFKGV 560  
DB 541 PPGNQEKDPLLKNQBFKGV 560

## RESULT 9

US-10-039-272A-2  
; Sequence 2, Application US/10039272A  
; Patent No. 6939955  
; GENERAL INFORMATION:  
; APPLICANT: RAMESHWAR, Pranela  
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE  
; FILE REFERENCE: 267/033 (UMD-0055)  
; CURRENT APPLICATION NUMBER: US/10/039,272A  
; CURRENT FILING DATE: 2001-10-20  
; PRIOR APPLICATION NUMBER: US 60/241,881  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.3  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-039-272A-2

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
DB 1 MECLYFGLFLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60  
QY 61 VMKGRDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
DB 61 VMKGRDMRWKNSWKGRRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPYNNWTAMSESDGNGTQSHHNPFGKPPHHPGWRWNFIYVFHTL 180  
DB 121 RNEAGLSADPYNNWTAMSESDGNGTQSHHNPFGKPPHHPGWRWNFIYVFHTL 180  
QY 181 GYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240  
DB 181 GYFQKLGRCVRSVNTANVTGLPOLMEVTVYRRHGRAYVPIAQKDVVYVTDQIPVFV 240  
QY 241 TFMQKNDNSDETFLKDLPIFMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TFMQKNDNSDETFLKDLPIFMFDVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLTAKAAAPGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
DB 301 HTYVLNGTFSNLTAKAAAPGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360

Qy 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPWPESSLLIDFVVTQCGSIPTVECTIIISDPT 420  
Db 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPWPESSLLIDFVVTQCGSIPTVECTIIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLISVDPDRPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLISVDPDRPAS 480  
Qy 481 PLRMANSLISVGLAIFVTVISLLVYKXKHEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db 481 PLRMANSLISVGLAIFVTVISLLVYKXKHEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Qy 541 PPGNQEKDPLLNQOEFGVS 560  
Db 541 PPGNQEKDPLLNQOEFGVS 560

RESULT 10  
US-10-007-700-225  
; Sequence 225, Application US/10007700  
; Patent No. 6960570  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedrick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Peckman, David W.  
; APPLICANT: Cai, Feng  
; APPLICANT: Foy, Teresa M.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C17  
; CURRENT APPLICATION NUMBER: US/10/007,700  
; CURRENT FILING DATE: 2001-11-30  
; NUMBER OF SEQ ID NOS: 469  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-007-700-225

Query Match 100.0%; Score 3026; DB 2; Length 560;  
Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKPHDVLGNERSAYNRHNLNGWSSDENDWNEKLYP 60  
Db 1 MECLYYFLGFLLLAARLPDAAKPHDVLGNERSAYNRHNLNGWSSDENDWNEKLYP 60  
Qy 61 VWKRGDMRWKNSWKGRRVQAVLTSDS PALVGSNITFAVNLIFPRCKEDANGNTVYEKNC 120  
Db 61 VWKRGDMRWKNSWKGRRVQAVLTSDS PALVGSNITFAVNLIFPRCKEDANGNTVYEKNC 120  
Qy 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHHPGWRWNFIYVFFHTL 180  
Db 121 RNEAGLSADPVYVNTAWSESDGNGTGQSHHNVFPDGKPPPHHPGWRWNFIYVFFHTL 180  
Qy 181 GQYFQKLGRCVRSVNTANTVLTGQLMEVTVYRRHGRAYVPIAQVKDVTQIIPVFV 240  
Db 181 GQYFQKLGRCVRSVNTANTVLTGQLMEVTVYRRHGRAYVPIAQVKDVTQIIPVFV 240  
Qy 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300

Db 241 TMFQKNDNRSSDETFLKDLPIFMDVLIHDPHFNLNSTYNTYKWSFGDNTGLFVSTNHTVN 300  
Qy 301 HTVYVNGTFSNLNTVKAAPGCPGPPPPRPSKPTPSLGPAGDNPLELSRIPDENCCOIN 360  
Db 301 HTVYVNGTFSNLNTVKAAPGCPGPPPPRPSKPTPSLGPAGDNPLELSRIPDENCCOIN 360  
Qy 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPWPESSLLIDFVVTQCGSIPTVECTIIISDPT 420  
Db 361 RYGHFOATITIVGILEVNIIOQTDVLMVPWPWPESSLLIDFVVTQCGSIPTVECTIIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLISVDPDRPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALSTLISVDPDRPAS 480  
Qy 481 PLRMANSLISVGLAIFVTVISLLVYKXKHEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Db 481 PLRMANSLISVGLAIFVTVISLLVYKXKHEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
Qy 541 PPGNQEKDPLLNQOEFGVS 560  
Db 541 PPGNQEKDPLLNQOEFGVS 560

RESULT 11  
US-09-985-799-90  
; Sequence 90, Application US/09985799  
; Patent No. RE38392  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/985,799  
; FILING DATE: 06-No. RE38392-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 90:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 90:

US-09-985-799-90

Query Match 100.0%; Score 3026; DB 5; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYVFGFLLAARLPDAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60  
 DB 1 MECLYVFGFLLAARLPDAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60

QY 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
 DB 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

QY 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFTL 180  
 DB 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFTL 180

QY 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYYVVTDOIIPVF 240  
 DB 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYYVVTDOIIPVF 240

QY 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300  
 DB 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300

QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360  
 DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360

QY 361 RYGHQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVECTIISDPT 420  
 DB 361 RYGHQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVECTIISDPT 420

QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480  
 DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480

QY 481 PLRMANSALISVGCIAIFVTVISLAVYKKGKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
 DB 481 PLRMANSALISVGCIAIFVTVISLAVYKKGKEYNPISPGNVVRSGKLSVFLNRAKAVF 540

RESULT 12

US-09-977-371-90  
 ; Sequence 90, Application US/09977371  
 ; Patent No. RE38490  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THOMPSON, Timothy C.  
 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
 ; NUMBER OF SEQUENCES: 175  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: BAKER & BOTTS, L.L.P.  
 ; STREET: 1299 Pennsylvania Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20004-2400  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq Version 1.5  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/977,371  
 ; FILING DATE: 16-Oct-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/594,031

FILING DATE: 30-JAN-1996  
 APPLICATION NUMBER: 60/006,838  
 FILING DATE: 16-NOV-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Remenick, James  
 REGISTRATION NUMBER: 36,902  
 REFERENCE/DOCKET NUMBER: 0A146-0110  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-639-7700  
 TELEFAX: 202-639-7890  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 90:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 560 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHEICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 SEQUENCE DESCRIPTION: SEQ ID NO: 90:

Query Match 100.0%; Score 3026; DB 5; Length 560;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-280;  
 Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYVFGFLLAARLPDAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60  
 DB 1 MECLYVFGFLLAARLPDAKRPDVLGNRPSPAYMREHNLNGSSDENDWNEKLYP 60

QY 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
 DB 61 VKRGDMRWKSKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120

QY 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFTL 180  
 DB 121 RNEAGLSADPYVYNNWTAWSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFTL 180

QY 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYYVVTDOIIPVF 240  
 DB 181 GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYYVVTDOIIPVF 240

QY 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300  
 DB 241 TMFQKNDNSSDETFKDLPIFMDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300

QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360  
 DB 301 HTYVLNGTFSNLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360

QY 361 RYGHQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVECTIISDPT 420  
 DB 361 RYGHQATITIVEGILEVNIQMTDVLMPVPPESSLDIFVVTCCGSIPTVECTIISDPT 420

QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480  
 DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLALTSTLISVPDRDPAS 480

QY 481 PLRMANSALISVGCIAIFVTVISLAVYKKGKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
 DB 481 PLRMANSALISVGCIAIFVTVISLAVYKKGKEYNPISPGNVVRSGKLSVFLNRAKAVF 540

RESULT 13

US-09-383-586-36  
 ; Sequence 36, Application US/09383586



Patent No. 6242419  
GENERAL INFORMATION:  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew  
APPLICANT: Abernethy, Nevin  
APPLICANT: Onrust, Rene  
APPLICANT: Murison, Greg  
TITLE OF INVENTION: Compounds isolated from stromal cells  
FILE REFERENCE: 11000.1037c1  
CURRENT APPLICATION NUMBER: US/09/383,586  
CURRENT FILING DATE: 1999-08-26  
NUMBER OF SEQ ID NOS: 38  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 36  
LENGTH: 574  
TYPE: PRT  
ORGANISM: Mouse  
US-09-383-586-36

Query Match 72.6%; Score 2197; DB 2; Length 574;  
Best Local Similarity 70.7%; Pred. No. 3.3e-201;  
Matches 403; Conservative 69; Mismatches 82; Indels 16; Gaps 2;

Qy 1 MECIYFLGFLLLAARLPDAAKRFHDVLGNRPSPAYMRHNLNGWSSDENDWNEKLYP 60  
Db 1 MESICGVLGFLLLAAGLPLOAKRFHDVLGHEQYPNHRHNLQURGHSSDENEDEHLYP 60

Qy 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITPAVNLIIPRCQKEDANGNIYVEKNC 120  
Db 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITPAVNLIIPRCQKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNTWTAWSEDSDENGTQSHHNVFPDGPFPHPGRWRNFIYVFHTL 180  
Db 121 RNDLGLTSDLHVYNTWTAGADGDWEDGTSRQHLRFPDRPFPRPHGKWKMSFYVVFHTL 180

Qy 181 GOYFQKLGRCRSVRVSVNTANTVLTGPQMEVTVYRRHGRVYPIAQVDVYVVTDIPIVFV 240  
Db 181 GOYFQKLGRCRSVRVSVNTANTVLTGPQMEVTVYRRHGRVYPIAQVDVYVVTDIPIVFV 240

Qy 241 TMSQKNDRLNLSDEIFLRLDPIVFDVLIHDPVSHFLNDSAISYKWNFGDNTGLFVSNHNTLN 300  
Db 241 TMSQKNDRLNLSDEIFLRLDPIVFDVLIHDPVSHFLNDSAISYKWNFGDNTGLFVSNHNTLN 300

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKPTPSLGPAGD 344  
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKPTPSLGPAGD 344

Qy 345 NPLELSRIPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWPSLSLIDFVVTVC 404  
Db 345 NPLELSRIPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWPSLSLIDFVVTVC 404

Qy 405 QGSIPTVECTIISDPTCEITQNTVCSVDVDEMCLLTVRRFTNGSGTYCVNLTGLGDDTSL 464  
Db 405 QGSIPTVECTIISDPTCEITQNTVCSVDVDEMCLLTVRRFTNGSGTYCVNLTGLGDDTSL 464

Qy 465 ALTSTLISVPRDRPASPLRMANSALISVGCCLAIPTVVISLIVYKHKHEYNPIENSPGNV 524  
Db 465 ALTSTLISVPRDRPASPLRMANSALISVGCCLAIPTVVISLIVYKHKHEYNPIENSPGNV 524

Qy 525 RSKGLSVFLNRAKAVFPFGNOEKDPLLNQ 554  
Db 541 KGKGLSVLLSHAKAPFFRGDQEKDPLLDQK 570

RESULT 14  
US-09-823-038A-36  
Sequence 36, Application US/09823038A  
Patent No. 6797271  
GENERAL INFORMATION:  
APPLICANT: Strachan, Lorna  
APPLICANT: Sleeman, Matthew

APPLICANT: Abernethy, Nevin  
APPLICANT: Onrust, Rene  
APPLICANT: Kumble, Anand  
APPLICANT: Murison, Greg  
TITLE OF INVENTION: Compositions Isolated From Stromal Cells  
FILE REFERENCE: 11000.1037c3  
CURRENT APPLICATION NUMBER: US/09/823,038A  
CURRENT FILING DATE: 2001-07-09  
NUMBER OF SEQ ID NOS: 61  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 36  
LENGTH: 574  
TYPE: PRT  
ORGANISM: Mouse  
US-09-823-038A-36

Query Match 72.6%; Score 2197; DB 2; Length 574;  
Best Local Similarity 70.7%; Pred. No. 3.3e-201;  
Matches 403; Conservative 69; Mismatches 82; Indels 16; Gaps 2;

Qy 1 MECIYFLGFLLLAARLPDAAKRFHDVLGNRPSPAYMRHNLNGWSSDENDWNEKLYP 60  
Db 1 MESICGVLGFLLLAAGLPLOAKRFHDVLGHEQYPNHRHNLQURGHSSDENEDEHLYP 60

Qy 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITPAVNLIIPRCQKEDANGNIYVEKNC 120  
Db 61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITPAVNLIIPRCQKEDANGNIYVEKNC 120

Qy 121 RNEAGLSADPYVYNTWTAWSEDSDENGTQSHHNVFPDGPFPHPGRWRNFIYVFHTL 180  
Db 121 RNDLGLTSDLHVYNTWTAGADGDWEDGTSRQHLRFPDRPFPRPHGKWKMSFYVVFHTL 180

Qy 181 GOYFQKLGRCRSVRVSVNTANTVLTGPQMEVTVYRRHGRVYPIAQVDVYVVTDIPIVFV 240  
Db 181 GOYFQKLGRCRSVRVSVNTANTVLTGPQMEVTVYRRHGRVYPIAQVDVYVVTDIPIVFV 240

Qy 241 TMSQKNDRLNLSDEIFLRLDPIVFDVLIHDPVSHFLNDSAISYKWNFGDNTGLFVSNHNTLN 300  
Db 241 TMSQKNDRLNLSDEIFLRLDPIVFDVLIHDPVSHFLNDSAISYKWNFGDNTGLFVSNHNTLN 300

Qy 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKPTPSLGPAGD 344  
Db 301 HTYVLNGTFSNLTVKAAAPGCP-----PPPPPRP-----SKPTPSLGPAGD 344

Qy 345 NPLELSRIPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWPSLSLIDFVVTVC 404  
Db 345 NPLELSRIPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVPWPSLSLIDFVVTVC 404

Qy 405 QGSIPTVECTIISDPTCEITQNTVCSVDVDEMCLLTVRRFTNGSGTYCVNLTGLGDDTSL 464  
Db 405 QGSIPTVECTIISDPTCEITQNTVCSVDVDEMCLLTVRRFTNGSGTYCVNLTGLGDDTSL 464

Qy 465 ALTSTLISVPRDRPASPLRMANSALISVGCCLAIPTVVISLIVYKHKHEYNPIENSPGNV 524  
Db 465 ALTSTLISVPRDRPASPLRMANSALISVGCCLAIPTVVISLIVYKHKHEYNPIENSPGNV 524

Qy 525 RSKGLSVFLNRAKAVFPFGNOEKDPLLNQ 554  
Db 541 KGKGLSVLLSHAKAPFFRGDQEKDPLLDQK 570

RESULT 15  
US-09-943-075A-5  
Sequence 5, Application US/09943075A  
Patent No. 6812002  
GENERAL INFORMATION:  
APPLICANT: Popoff, Steven N.  
APPLICANT: Safado, Faye F.  
APPLICANT: Owen, Thomas A.  
APPLICANT: Smock, Steven L.  
TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,  
; TITLE OF INVENTION: Osteoactivin Protein and Methods of Stimulating Bone Differentiation



```
FILE REFERENCE: 71369.262
CURRENT APPLICATION NUMBER: US/09/943,075A
CURRENT FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: US 60/229,006
PRIOR FILING DATE: 2000-08-30
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 574
TYPE: PRT
ORGANISM: Mouse
US-09-943-075A-5

Query Match      72.6%; Score 2197; DB 2; Length 574;
Best Local Similarity 70.7%; Pred. No. 3.3e-201;
Matches 403; Conservative 69; Mismatches 82; Indels 16; Gaps 2;

QY      1 MESLYYFGLFLLAARLPLDAARFHDVILGNRPASVMEHNLQNGWSSDENDWNEKLYP 60
DB      1 MESLCVGLFLLAAGLPLOAAKRFDFVLGHEQYDPHMEHNLQNGWSSDENEDEHLYP 60

QY      61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITPAVNLIFPRCQKEDANGNIYYEKNC 120
DB      61 VWRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITPAVNLIFPRCQKEDANGNIYYEKNC 120

QY      121 RNEAGLSADPYVNTWASEDSGNGTCQSHHNVPDGPDPHPGWRNPNPIYVFFHTL 180
DB      121 RNDLGLTSLHVNNTAGADGDWEDGTSRQHLRFPDRRPPFRPHGWMKNSFVYVFFHTL 180

QY      181 GQYFQKLGRCVSVRVNTANVTLPQLMETVYVRHGRAYVPIAOKVYVYVTDQIPVFV 240
DB      181 GQYFQKLGRCARSINTVNLTAGQVMEVTVFRYGRAYIIFISKVKVIVITDQIPVFV 240

QY      241 TMFQKNDNRSSDETFKOLPIMFDVLIHDPHFHLYSTINYKWSFGDNTGLEFVSTNHTVN 300
DB      241 TWSQKNDNRNLSDEIFLRLPIVPDVLHDPHFHLYSDAISYKWNFGDNTGLEFVSNHTLN 300

QY      301 HTYVLNGTFSLNLTWKAAAPGCPPPPP--PPRP-----SKPTPSLGPAGD 344
DB      301 HTYVLNGTFNLNLTQAVPGCPPPSPSTPPSPSTPPSPPLTSTPSPLMPTGY 360

QY      345 NPLELSRIPDENCOINRYGHFOATITVEGILEVNIQMTDVLMPVWPBESSLIDFVTC 404
DB      361 KSMELSDISNENCRINRYGYFRATITVEGILEVSIQIADVPMPPTQPANSLMDFTVTC 420

QY      405 QGSIPTVECTIISDPTCEITQNTVCSPVDVDEMCLLTVRTFENGSGTYCVNLTGDDTSL 464
DB      421 KGATPWEACTIISDPTCQIAQNRVCSVAVDGLCLLSVRAENGSGTYCVNLTGDDASL 480

QY      465 ALTSTLISVPDRDPASPLRMANSALISVGCLAI FVTVISLLVYKXKHEYNPIENSPGNV 524
DB      481 ALTSTLISIPGKDPSPRAVNGVLISIGCLAVLVTMTVILLYKXKAYKPIGNCPRNTV 540

QY      525 RSKGLSVFLNRAKAVFPFGNQEKKDPLLNQ 554
DB      541 KKGGLSVLLSHAKAPFPRGDEKQDPLQDK 570
```

Search completed: December 24, 2005, 04:37:24  
Job time : 52 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	3026	100.0	560	3	US-09-735-705-225	Sequence 225, App
2	3026	100.0	560	3	US-09-850-718A-225	Sequence 225, App
3	3026	100.0	560	3	US-09-897-778-225	Sequence 225, App
4	3026	100.0	560	3	US-09-943-075A-6	Sequence 6, Appli
5	3026	100.0	560	4	US-10-039-272-2	Sequence 2, Appli
6	3026	100.0	560	4	US-10-007-700-225	Sequence 225, App
7	3026	100.0	560	4	US-10-117-982-225	Sequence 225, App
8	3026	100.0	560	4	US-10-463-106-2	Sequence 2, Appli
9	3026	100.0	560	4	US-10-295-027-1258	Sequence 1258, Ap
10	3026	100.0	560	4	US-10-313-986-225	Sequence 225, App
11	3026	100.0	560	4	US-10-309-290-152	Sequence 152, App
12	3026	100.0	560	4	US-10-408-765A-466	Sequence 466, App
13	3026	100.0	560	5	US-10-775-972-225	Sequence 225, App
14	3026	100.0	560	5	US-10-723-860-1128	Sequence 1128, Ap
15	3026	100.0	560	5	US-10-922-124-225	Sequence 225, App
16	3026	100.0	560	5	US-10-978-758-6	Sequence 6, Appli
17	3026	100.0	560	5	US-10-631-467-821	Sequence 821, App
18	3006	99.3	572	4	US-10-227-884-42	Sequence 42, Appl
19	3006	99.3	572	4	US-10-230-163-42	Sequence 42, Appl
20	3006	99.3	572	4	US-10-230-338-42	Sequence 42, Appl
21	3006	99.3	572	4	US-10-218-631-42	Sequence 42, Appl
22	3006	99.3	572	4	US-10-230-414-42	Sequence 42, Appl
23	3006	99.3	572	4	US-10-232-224-42	Sequence 42, Appl
24	3006	99.3	572	4	US-10-216-159A-42	Sequence 42, Appl
25	3006	99.3	572	4	US-10-218-849-42	Sequence 42, Appl
26	3006	99.3	572	4	US-10-227-873-42	Sequence 42, Appl
27	3006	99.3	572	4	US-10-227-883-42	Sequence 42, Appl

```
Qy 241 TMFQKDRNSDSETFKDLPLIMFDVLHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKDRNSDSETFKDLPLIMFDVLHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYYVLNGTFSNLNLTAKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Db 301 HTYYVLNGTFSNLNLTAKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Qy 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCNVNLTLGDDTSLALTSTLISVDPDRPAS 480
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCNVNLTLGDDTSLALTSTLISVDPDRPAS 480
Qy 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNKQEFKGV 560
Db 541 FPGNQEKDPLLNKQEFKGV 560
```

## RESULT 2

```
US-09-850-716A-225
; Sequence 225, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retex, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-225
```

```
Query Match 100.0%; Score 3026; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGLFLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Db 1 MECLYYFLGLFLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Qy 61 VWKRGDMRWKNSWKGRRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVYKNC 120
Db 61 VWKRGDMRWKNSWKGRRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVYKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQSHNVFPDGKFPFPHPGWRRWNFIYVFHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQSHNVFPDGKFPFPHPGWRRWNFIYVFHTL 180
Qy 181 GOYFQKLGRCRSVRVSNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPVFV 240
Db 181 GOYFQKLGRCRSVRVSNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPVFV 240
Qy 241 TMFQKDRNSDSETFKDLPLIMFDVLHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKDRNSDSETFKDLPLIMFDVLHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYYVLNGTFSNLNLTAKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Db 301 HTYYVLNGTFSNLNLTAKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
```

```
Qy 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Db 361 RYGHFOATITIVGILEVNIIOQMTDVLMPVPWPPESSLIDFVTCQGSIPTEVCTIISDPT 420
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCNVNLTLGDDTSLALTSTLISVDPDRPAS 480
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCNVNLTLGDDTSLALTSTLISVDPDRPAS 480
Qy 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Db 481 PLRWANSALISVGCIAIFVTVVISLLVYKHKHKEYNPNIENSPGNVVRSGLSVFLNRAKAVF 540
Qy 541 FPGNQEKDPLLNKQEFKGV 560
Db 541 FPGNQEKDPLLNKQEFKGV 560

RESULT 3
US-09-897-778-225
; Sequence 225, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darriack
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-225
```

```
Query Match 100.0%; Score 3026; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGLFLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Db 1 MECLYYFLGLFLLAARLPDAAKRFHDVLGNRPSAYMRHNLNGWSSDENDWNEKLYP 60
Qy 61 VWKRGDMRWKNSWKGRRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVYKNC 120
Db 61 VWKRGDMRWKNSWKGRRVQAVLTSDSPALVGSNTITFAVNLIFPRCKEDANGNIYVYKNC 120
Qy 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQSHNVFPDGKFPFPHPGWRRWNFIYVFHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQSHNVFPDGKFPFPHPGWRRWNFIYVFHTL 180
Qy 181 GOYFQKLGRCRSVRVSNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPVFV 240
Db 181 GOYFQKLGRCRSVRVSNTANTLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTQIIPVFV 240
Qy 241 TMFQKDRNSDSETFKDLPLIMFDVLHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKDRNSDSETFKDLPLIMFDVLHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Qy 301 HTYYVLNGTFSNLNLTAKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
Db 301 HTYYVLNGTFSNLNLTAKAAAPGCPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCOIN 360
```

QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420  
DB 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420  
QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPDRDPAS 480  
DB 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPDRDPAS 480  
QY 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
QY 541 PFGNQEKDPLLNQKQEFKGS 560  
DB 541 PFGNQEKDPLLNQKQEFKGS 560

## RESULT 4

US-09-943-075A-6  
; Sequence 6, Application US/09943075A  
; Patent No. US20020151486A1  
; GENERAL INFORMATION:  
; APPLICANT: Popoff, Steven N.  
; APPLICANT: Safado, Favez F.  
; APPLICANT: Owen, Thomas A.  
; APPLICANT: Smock, Steven L.  
; TITLE OF INVENTION: Osteocalcin Protein and Nucleic Acids Encoding the Same,  
; TITLE OF INVENTION: Compositions and Methods of Stimulating Bone Differentiation  
; FILE REFERENCE: 71369.262  
; CURRENT APPLICATION NUMBER: US/09/943,075A  
; PRIOR FILING DATE: 2001-08-30  
; PRIOR APPLICATION NUMBER: US 60/229,006  
; PRIOR FILING DATE: 2000-08-30  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Human  
US-09-943-075A-6

Query Match 100.0%; Score 3026; DB 3; Length 560;  
Best Local Similarity 100.0%; Pred. No. 9e-246;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFFLGFLLLAARLPDAAKRPDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60  
DB 1 MECLYFFLGFLLLAARLPDAAKRPDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60  
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGPHPHGRWRNFIYVFHTL 180  
DB 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGPHPHGRWRNFIYVFHTL 180  
QY 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYPIAQVKDYYVVTQIPVFV 240  
DB 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYPIAQVKDYYVVTQIPVFV 240  
QY 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLVTAAAGPCPPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360  
DB 301 HTYVLNGTFSNLVTAAAGPCPPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360  
QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420  
DB 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420

QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPDRDPAS 480  
DB 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPDRDPAS 480  
QY 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
QY 541 PFGNQEKDPLLNQKQEFKGS 560  
DB 541 PFGNQEKDPLLNQKQEFKGS 560

## RESULT 5

US-10-039-272-2  
; Sequence 2, Application US/10039272  
; Publication No. US20020168653A1  
; GENERAL INFORMATION:  
; APPLICANT: RAMESHWAR, Pranela  
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE  
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey  
; CURRENT APPLICATION NUMBER: US/10/039,272  
; CURRENT FILING DATE: 2001-10-20  
; PRIOR APPLICATION NUMBER: US 60/241,881  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-039-272-2

Query Match 100.0%; Score 3026; DB 4; Length 560;  
Best Local Similarity 100.0%; Pred. No. 9e-246;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFFLGFLLLAARLPDAAKRPDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60  
DB 1 MECLYFFLGFLLLAARLPDAAKRPDVLGNERPSAYMRHNLQNGWSSDENDWNEKLYP 60  
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
DB 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGPHPHGRWRNFIYVFHTL 180  
DB 121 RNEAGLSADPVYVNTAWSESDGNGTQSHHNVFPDGPHPHGRWRNFIYVFHTL 180  
QY 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYPIAQVKDYYVVTQIPVFV 240  
DB 181 QOYFQKLGRCVRSVNTANTVLGQPMVEVTVYRRHGRAYPIAQVKDYYVVTQIPVFV 240  
QY 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TMFQKNDNSDETFLKDLPIFMFVLIHDPHSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTFSNLVTAAAGPCPPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360  
DB 301 HTYVLNGTFSNLVTAAAGPCPPPPPPRPSKPTPSLGPAGDNPLSLRIPDENCQIN 360  
QY 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420  
DB 361 RYGHFOATITIVGILEVNIQMTDVLMPVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420  
QY 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPDRDPAS 480  
DB 421 CEITONTVCSPVDVDEMCLLTVRRTFNGSGTYCVNLTIGDDTSLALTSTLISVPDRDPAS 480  
QY 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPISPGNVVRSGKLSVFLNRAKAVF 540  
DB 481 PLRMANSALISVGCCLAIPTVVISLVYKKHKEYNPISPGNVVRSGKLSVFLNRAKAVF 540

```
QY 541 FPGNQEKDPLLNQEFKGV 560
Db 541 FPGNQEKDPLLNQEFKGV 560

RESULT 6
US-10-007-700-225
; Sequence 225, Application US/10007700
; Publication No. US20030064947A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Li, Samuel X.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007.700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-225

Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFHTL 180
QY 181 GOYFQKLGRCRSVRVSVNTANVTLPQLMVTVYRRHGRAYVPIAQKDVVVTDDQIIPFV 240
Db 181 GOYFQKLGRCRSVRVSVNTANVTLPQLMVTVYRRHGRAYVPIAQKDVVVTDDQIIPFV 240
QY 241 TMFQKNDNRNSDDETFKDLPIMFVDLHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRNSDDETFKDLPIMFVDLHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLSLRIPDENQCIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLSLRIPDENQCIN 360
QY 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420
Db 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSIALTSLTISVDRDPAS 480

RESULT 7
US-10-117-982-225
; Sequence 225, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara
; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117.982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-225

Query Match 100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60
Db 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSPAYMRHNLQNGSSDENDWNEKLYP 60
QY 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIFPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFHTL 180
Db 121 RNEAGLSADPYVYNWTAWSESDSGENGTSQHSHNVFPDGKPPHPGWRWNFIYVFHTL 180
QY 181 GOYFQKLGRCRSVRVSVNTANVTLPQLMVTVYRRHGRAYVPIAQKDVVVTDDQIIPFV 240
Db 181 GOYFQKLGRCRSVRVSVNTANVTLPQLMVTVYRRHGRAYVPIAQKDVVVTDDQIIPFV 240
QY 241 TMFQKNDNRNSDDETFKDLPIMFVDLHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
Db 241 TMFQKNDNRNSDDETFKDLPIMFVDLHDPSSHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLSLRIPDENQCIN 360
Db 301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPSPKPTPSLGPAGNPLSLRIPDENQCIN 360
QY 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420
Db 361 RYGHFQATITIVGILEVNIQMTDVLMPVPWPESSLIIDFVTCQGSIPTEVCTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGGDTSIALTSLTISVDRDPAS 480
```

QY 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLATSTLISVPDRPAS 480  
DB 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLATSTLISVPDRPAS 480  
QY 481 PLRMANALISVGLCLAFVTVISLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
DB 481 PLRMANALISVGLCLAFVTVISLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
QY 541 PFGNQEKDPLLKNQBFKGV 560  
DB 541 PFGNQEKDPLLKNQBFKGV 560  
RESULT 8  
US-10-463-106-2  
; Sequence 2, Application US/10463106  
; Publication No. US20030202938A1  
; GENERAL INFORMATION:  
; APPLICANT: RAMESHWAR, Pranela  
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-1 GENE  
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey  
; CURRENT APPLICATION NUMBER: US/10/463,106  
; CURRENT FILING DATE: 2003-06-17  
; PRIOR APPLICATION NUMBER: US 10/039,272  
; PRIOR FILING DATE: 2001-10-20  
; PRIOR APPLICATION NUMBER: US 60/241,881  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-463-106-2  
Query Match 100.0%; Score 3026; DB 4; Length 560;  
Best Local Similarity 100.0%; Pred. No. 9e-246;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60  
DB 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60  
QY 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
DB 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPVYNNWTAMSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180  
DB 121 RNEAGLSADPVYNNWTAMSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180  
QY 181 GOYFQKGRCSVRVSVNTANVTGLPQLMEVTVYRRHGRAYVPIAQVDVYVVDQIPVFV 240  
DB 181 GOYFQKGRCSVRVSVNTANVTGLPQLMEVTVYRRHGRAYVPIAQVDVYVVDQIPVFV 240  
QY 241 TMFQKDRNNSDETFLKPLIMFVDLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
DB 241 TMFQKDRNNSDETFLKPLIMFVDLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300  
QY 301 HTYVLNGTSLNLTUKAAGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
DB 301 HTYVLNGTSLNLTUKAAGCPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDPDP 360  
QY 361 RYGHFQATITVEGILEVNIQMTDVLMPVPESLIDFVVTCCSGIPTVECTIISDPT 420  
DB 361 RYGHFQATITVEGILEVNIQMTDVLMPVPESLIDFVVTCCSGIPTVECTIISDPT 420  
QY 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLATSTLISVPDRPAS 480  
DB 421 CEITONTVCSPVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLATSTLISVPDRPAS 480  
QY 481 PLRMANALISVGLCLAFVTVISLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540

DB 481 PLRMANALISVGLCLAFVTVISLVYKHKKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540  
QY 541 PFGNQEKDPLLKNQBFKGV 560  
DB 541 PFGNQEKDPLLKNQBFKGV 560  
RESULT 9  
US-10-295-027-1258  
; Sequence 1258, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevezi, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
; FILE REFERENCE: 018501-012500US  
; CURRENT APPLICATION NUMBER: US/10/295,027  
; CURRENT FILING DATE: 2002-11-13  
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1258  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-295-027-1258  
Query Match 100.0%; Score 3026; DB 4; Length 560;  
Best Local Similarity 100.0%; Pred. No. 9e-246;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60  
DB 1 MECLYFLGFLLLAARLPDAAKRFHDLGNRPSAYMREHNLQNGSSDENDWNEKLYP 60  
QY 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
DB 61 VKRGDMWKNWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120  
QY 121 RNEAGLSADPVYNNWTAMSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180  
DB 121 RNEAGLSADPVYNNWTAMSESDGNGTGQSHHNVFPDGKPPHPGWRWNFIYVFHTL 180  
QY 181 GOYFQKGRCSVRVSVNTANVTGLPQLMEVTVYRRHGRAYVPIAQVDVYVVDQIPVFV 240

Db 181 GQYFQKLGRCVRSVNTANVTGLGQMEVTVYRRHGRAYVPIAQVXDVVVTQIEVFV 240  
Qy 241 TMFQKNDNRNSDDEFLKDLPIIMFDVLHDPGSHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300  
Db 241 TMFQKNDNRNSDDEFLKDLPIIMFDVLHDPGSHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300  
Qy 301 HTYYLNGTFTSLNLTVAAGPCPPPPPPRPSKPTPSLGPAGDNPLLELSRIPDENQCIN 360  
Db 301 HTYYLNGTFTSLNLTVAAGPCPPPPPPRPSKPTPSLGPAGDNPLLELSRIPDENQCIN 360  
Qy 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQSGIPTVEVCTIISDPT 420  
Db 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQSGIPTVEVCTIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480  
Qy 481 PLRWANSALISVGCLAIFFVTVISLLVYKHKHKEYNPIENSPGNVVRSGKLSVFLNRAKAVF 540  
Db 481 PLRWANSALISVGCLAIFFVTVISLLVYKHKHKEYNPIENSPGNVVRSGKLSVFLNRAKAVF 540  
Qy 541 FPGNQEKDPLLNQOEKFGVGS 560  
Db 541 FPGNQEKDPLLNQOEKFGVGS 560  
RESULT 10  
US-10-313-986-225  
; Sequence 225, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Tereesa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313.986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-313-986-225  
Query Match 100.0%; Score 3026; DB 4; Length 560;  
Best Local Similarity 100.0%; Pred. No. 9e-246;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MECLYYFLGFLLLAARLPDAAKGFHDVLGNERSAYMRHNLNGHSSDENDWNEKLYP 60  
Db 1 MECLYYFLGFLLLAARLPDAAKGFHDVLGNERSAYMRHNLNGHSSDENDWNEKLYP 60  
Qy 61 VWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTTFAVNLIFPRCKEDANGNIYBKNC 120  
Db 61 VWKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTTFAVNLIFPRCKEDANGNIYBKNC 120  
Qy 121 RNEAGLSADPVYVNTWTAWSESDGNGTGGSHNVFPDGKPFPHHPCGRRWNFIYVFHTL 180  
Db 121 RNEAGLSADPVYVNTWTAWSESDGNGTGGSHNVFPDGKPFPHHPCGRRWNFIYVFHTL 180  
Qy 181 GQYFQKLGRCVRSVNTANVTGLGQMEVTVYRRHGRAYVPIAQVXDVVVTQIEVFV 240  
Db 181 GQYFQKLGRCVRSVNTANVTGLGQMEVTVYRRHGRAYVPIAQVXDVVVTQIEVFV 240  
Qy 241 TMFQKNDNRNSDDEFLKDLPIIMFDVLHDPGSHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300

Db 241 TMFQKNDNRNSDDEFLKDLPIIMFDVLHDPGSHFLNVTINYNKWSFGDNTGLFVSTNHTVN 300  
Qy 301 HTYYLNGTFTSLNLTVAAGPCPPPPPPRPSKPTPSLGPAGDNPLLELSRIPDENQCIN 360  
Db 301 HTYYLNGTFTSLNLTVAAGPCPPPPPPRPSKPTPSLGPAGDNPLLELSRIPDENQCIN 360  
Qy 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQSGIPTVEVCTIISDPT 420  
Db 361 RYGHFQATITIVBGEILEVNIIOQMTDVLMPVPWPPESSLIDFVVTCCQSGIPTVEVCTIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTGLDDTSLALTSTLISVDPDRPAS 480  
Qy 481 PLRWANSALISVGCLAIFFVTVISLLVYKHKHKEYNPIENSPGNVVRSGKLSVFLNRAKAVF 540  
Db 481 PLRWANSALISVGCLAIFFVTVISLLVYKHKHKEYNPIENSPGNVVRSGKLSVFLNRAKAVF 540  
Qy 541 FPGNQEKDPLLNQOEKFGVGS 560  
Db 541 FPGNQEKDPLLNQOEKFGVGS 560  
RESULT 11  
US-10-309-290-152  
; Sequence 152, Application US/10309290  
; Publication No. US20040023241A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsbrook II, John P.  
; APPLICANT: Anderson, David W.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Chillakuru, Rajeev A.  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Gerlach, Valerie L.  
; APPLICANT: Gorman, Linda  
; APPLICANT: Gould-Rothberg, Bonnie E.  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Jeffers, Michael E.  
; APPLICANT: Ji, Weizhen  
; APPLICANT: Li, Li  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Murphey, Ryan  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Peyman, John A.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Smithson, Glennda  
; APPLICANT: Starling, Gary  
; APPLICANT: Taupier, Raymond J.  
; APPLICANT: Voess, Edward Z.  
; APPLICANT: Zhong, Haihong  
; APPLICANT: Zhong, Mei  
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
; FILE REFERENCE: 21402-502A  
; CURRENT APPLICATION NUMBER: US/10/309,290  
; CURRENT FILING DATE: 2002-12-02  
; PRIOR APPLICATION NUMBER: 60/336,600  
; PRIOR FILING DATE: 2001-12-05  
; PRIOR APPLICATION NUMBER: 60/338,285  
; PRIOR FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/341,346  
; PRIOR FILING DATE: 2001-12-12  
; PRIOR APPLICATION NUMBER: 60/341,477  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: 60/341,540  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: 60/342,592  
; PRIOR FILING DATE: 2001-12-20  
; PRIOR APPLICATION NUMBER: 60/344,297  
; PRIOR FILING DATE: 2001-12-27



```
; PRIORITY APPLICATION NUMBER: 60/344,903
; PRIORITY FILING DATE: 2001-12-31
; PRIORITY APPLICATION NUMBER: 60/373,288
; PRIORITY FILING DATE: 2002-04-17
; PRIORITY APPLICATION NUMBER: 60/380,981
; PRIORITY FILING DATE: 2002-05-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 152
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-309-290-152

Query Match      100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAARFHDVILGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYFGLFLLAARLPDAARFHDVILGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
DB 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
DB 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
QY 181 QGYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIIPFV 240
DB 181 QGYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIIPFV 240
QY 241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPSPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPSPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLSLSRIPDENCQIN 360
DB 301 HTYVLNGTFSNLTKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLSLSRIPDENCQIN 360
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPSSLDIFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPSSLDIFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
QY 481 PLRMANSALISVGLCIAIFVTVISLLVYKKHKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGLCIAIFVTVISLLVYKKHKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQEFKGV 560
DB 541 PFGNQEKDPLLKNQEFKGV 560

RESULT 12
US-10-408-765A-466
; Sequence 466, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
```

```
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408.765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 466
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-466

Query Match      100.0%; Score 3026; DB 4; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MECLYFGLFLLAARLPDAARFHDVILGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
DB 1 MECLYFGLFLLAARLPDAARFHDVILGNRPSPAYMRHNLQNGWSSDENDWNEKLYP 60
QY 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
DB 61 VVKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYVEKNC 120
QY 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
DB 121 RNEAGLSADPYVYNNWTAMSESDGNGTGQSHHNVPDPKPPHHPGRRWNFIYVFTL 180
QY 181 QGYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIIPFV 240
DB 181 QGYFQKLGRCVRSVNTANTVTLGPQLMEVTVYRRHGRAYVPIAQKDVYVVDQIIPFV 240
QY 241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPSPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
DB 241 TMFQKNDNRSSDETFLKOLPIMFDVLIHDPSPHFLNYSTINYKWSFGDNTGLFVSTNHTVN 300
QY 301 HTYVLNGTFSNLTKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLSLSRIPDENCQIN 360
DB 301 HTYVLNGTFSNLTKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLSLSRIPDENCQIN 360
QY 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPSSLDIFVVTCCGSIPTVECTIISDPT 420
DB 361 RYGHFOATITIVEGILEVNIQMTDVLMPVPWPSSLDIFVVTCCGSIPTVECTIISDPT 420
QY 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
DB 421 CEITQNTVCSVDVDEMCLLTVRRTFNGSGTYCVNLTLGDDTSLATSTLISVPDRDPAS 480
QY 481 PLRMANSALISVGLCIAIFVTVISLLVYKKHKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
DB 481 PLRMANSALISVGLCIAIFVTVISLLVYKKHKEYNPIENSPGNVRSKGLSVFLNRAKAVF 540
QY 541 PFGNQEKDPLLKNQEFKGV 560
DB 541 PFGNQEKDPLLKNQEFKGV 560

RESULT 13
US-10-775-972-225
; Sequence 225, Application US/10775972
; Publication No. US20040235072A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C21
; CURRENT APPLICATION NUMBER: US/10/775,972
; CURRENT FILING DATE: 2004-02-10
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
```



```
; PRIOR APPLICATION NUMBER: 09/685,696
; PRIOR FILING DATE: 2000-10-09
; PRIOR APPLICATION NUMBER: 09/662,786
; PRIOR FILING DATE: 2000-09-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 563
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-922-124-225

Query Match      100.0%; Score 3026; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 9e-246;
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MECLYFGLGLLAARLPDAAKRFHDVLGNRPSGAYMREHNLQNGWSSDENDWNEKLYP 60
Db      1 MECLYFGLGLLAARLPDAAKRFHDVLGNRPSGAYMREHNLQNGWSSDENDWNEKLYP 60

QY      61 VMKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120
Db      61 VMKRGDMRWKNSKGRVQAVLTSDSPALVGSNITFAVNLIFFRCQKEDANGNIYVEKNC 120

QY      121 RNEAGLSADPYNTNTANSESDGNGTQSHHNVPDGGKPPHHPGWRWNFIYVFTL 180
Db      121 RNEAGLSADPYNTNTANSESDGNGTQSHHNVPDGGKPPHHPGWRWNFIYVFTL 180

QY      181 GOYFQKLGRCVRSVNTANVTLGQLMETVYRRHGRAYVPIAQVKVYVVTDOIPIVFV 240
Db      181 GOYFQKLGRCVRSVNTANVTLGQLMETVYRRHGRAYVPIAQVKVYVVTDOIPIVFV 240

QY      241 TMFQKNDNRSSDETFKDLPIIMFDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300
Db      241 TMFQKNDNRSSDETFKDLPIIMFDVLIHDPHFNLNSTINYKWSFGDNTGLFVSTNHTVN 300

QY      301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360
Db      301 HTYVLNGTFSNLTVKAAAPGCPPPPPPPRPSKPTPSLGPAGDNPLELSRIPDENCQIN 360

QY      361 RYGHFQATITIVEGILEVNIIOQTDVLMVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420
Db      361 RYGHFQATITIVEGILEVNIIOQTDVLMVPWPPESSLIDFVVTCCGSIPTVECTIISDPT 420

QY      421 CEITQNTVCSVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLALTSTLISVPDRDPAS 480
Db      421 CEITQNTVCSVDVDEMCLLTVRRFTNGSGTTCVNLTLGDDTSLALTSTLISVPDRDPAS 480

QY      481 PLRMANSALISVGLAIFVTVISLLVYKHKYKYNPIENSPGNVVRSGLSVFLNRAKAVF 540
Db      481 PLRMANSALISVGLAIFVTVISLLVYKHKYKYNPIENSPGNVVRSGLSVFLNRAKAVF 540

QY      541 PFGNQEKDPLLKNQEPKGV 560
Db      541 PFGNQEKDPLLKNQEPKGV 560
```

Search completed: December 24, 2005, 04:49:56  
Job time : 166 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 24, 2005, 04:36:36 ; Search time 13 Seconds  
(without alignments)  
307.260 Million cell updates/sec

Title: US-09-853-880A-17  
Perfect score: 3026  
Sequence: 1 MECLYFLGLLLAARPLD.....PFGNQEKDPLLNQEPKQVS 560

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 54001 seqs, 7132810 residues

Total number of hits satisfying chosen parameters: 54001

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New:  
1: /cgm2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgm2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
3: /cgm2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
4: /cgm2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
5: /cgm2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
6: /cgm2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
7: /cgm2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
8: /cgm2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3026	100.0	560	US-10-623-155-225	Sequence 225, App
2	3026	100.0	563	US-10-821-234-1067	Sequence 1067, Ap
3	1029	34.0	206	US-11-090-439-44	Sequence 44, Appl
4	622.5	20.6	662	US-11-090-439-9	Sequence 9, Appl
5	107.5	3.6	564	US-11-022-289-10	Sequence 10, Appl
6	105.5	3.5	427	US-10-995-561-816	Sequence 816, App
7	105.5	3.5	427	US-10-995-561-818	Sequence 818, App
8	105.5	3.5	452	US-10-995-561-817	Sequence 817, App
9	97	3.2	1717	US-11-192-967-2	Sequence 2, Appl
10	97	3.2	1717	US-11-193-715-2	Sequence 2, Appl
11	96.5	3.2	583	US-11-080-991-64	Sequence 64, Appl
12	94.5	3.1	557	US-11-022-289-6	Sequence 6, Appl
13	92	3.0	432	US-11-140-417-4	Sequence 4, Appl
14	92	3.0	435	US-10-510-386-62	Sequence 62, Appl
15	90.5	3.0	551	US-11-022-289-7	Sequence 7, Appl
16	90.5	3.0	557	US-11-022-289-2	Sequence 2, Appl
17	90	3.0	610	US-10-606-302-1	Sequence 1, Appl
18	89.5	3.0	438	US-11-140-417-2	Sequence 2, Appl
19	89.5	3.0	551	US-11-022-289-8	Sequence 8, Appl
20	87.5	2.9	557	US-11-022-289-5	Sequence 5, Appl
21	86.5	2.9	743	US-10-485-517-351	Sequence 351, App
22	86.5	2.9	877	US-10-485-517-200	Sequence 200, App
23	86	2.8	626	US-11-010-748A-1	Sequence 1, Appl
24	85.5	2.8	1390	US-11-063-343-35	Sequence 35, Appl
25	84.5	2.8	175	US-10-821-234-1074	Sequence 1074, Ap

ALIGNMENTS

RESULT 1

US-10-623-155-225  
; Sequence 225, Application US/10623155  
; Publication No. US20050261166A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Peckham, David W.  
; APPLICANT: Retter, Gary W.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C20  
; CURRENT APPLICATION NUMBER: US/10/623,155  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-623-155-225

Query Match	100.0%	Score 3026;	DB 6;	Length 560;
Best Local Similarity	100.0%;	Pred. No. 4.2e-255;	Mismatches 0;	Indels 0; Gaps 0;
Matches 560;	Conservative 0;			
QY	1	MECLYFLGLLLAARPLDAAKRFHDVIGNERPSAYMREHNQNGSSDENDWNEKLYP	60	
Db	1	MECLYFLGLLLAARPLDAAKRFHDVIGNERPSAYMREHNQNGSSDENDWNEKLYP	60	
QY	61	VWKGDMWKNKSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC	120	
Db	61	VWKGDMWKNKSWKGRVQAVLTSDSPALVGSNITFAVNLIPPRCKEDANGNIYKNC	120	
QY	121	RNEAGLSADPYVYNWNTAWSESDGNGTQSHHNVFPDGKPPHPGWRWNIYVFHTL	180	
Db	121	RNEAGLSADPYVYNWNTAWSESDGNGTQSHHNVFPDGKPPHPGWRWNIYVFHTL	180	
QY	181	GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVYVTDQIPVF	240	
Db	181	GOYFQKGRCSVRVSVNTANVTGLPQMEVTVYRHRGRAYVPIAQVKDYVYVTDQIPVF	240	
QY	241	TNFKQNRNNSDETFLKDLPIPMFVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN	300	
Db	241	TNFKQNRNNSDETFLKDLPIPMFVLIHDPHFLNYSTINYKWSFGDNTGLFVSTNHTVN	300	
QY	301	HTYVNLNGTFSNLVTYKAAAGPCPPPPPPPSKPTPSIGPAGDNPLEISRPDENCOIN	360	

Db 301 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTSLGPAGDNPLLSRIPDENCQIN 360  
Qy 361 RYGHFOATTIVGILEVNIIOQMTDVLMPWPBPSSLIIDFVVTQCGSIPTEVCTIIISDPT 420  
Db 361 RYGHFOATTIVGILEVNIIOQMTDVLMPWPBPSSLIIDFVVTQCGSIPTEVCTIIISDPT 420  
Qy 421 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALSTLSVDRDRPAS 480  
Db 421 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALSTLSVDRDRPAS 480  
Qy 481 PLRMANALISVGCLAFVTVVISLLVYKHKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540  
Db 481 PLRMANALISVGCLAFVTVVISLLVYKHKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540  
Qy 541 FPGNQEKDPLLNQEPKGV 560  
Db 541 FPGNQEKDPLLNQEPKGV 560

RESULT 2  
US-10-821-234-1067  
; Sequence 1067, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; FILE REFERENCE: 821A  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 1067  
; LENGTH: 563  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-821-234-1067

Query Match 100.0%; Score 3026; DB 6; Length 563;  
Best Local Similarity 100.0%; Pred. No. 4.2e-255;  
Matches 560; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGWSSDENDWNEKLYP 60  
Db 4 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGWSSDENDWNEKLYP 63  
Qy 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIIPRCOKEDANGNIYVEKNC 120  
Db 64 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIIPRCOKEDANGNIYVEKNC 123  
Qy 121 RNEAGLSADPVYVNTWTAWSESDGNGTGSQSHNVFPDGPFPHPGRRWNFIYVPHTL 180  
Db 124 RNEAGLSADPVYVNTWTAWSESDGNGTGSQSHNVFPDGPFPHPGRRWNFIYVPHTL 183  
Qy 181 GOYFQKLGRCSVRVSVNTANTVLGQPMVTVYRRHGRVYVYIAQVXDYVVTDIQIPVF 240  
Db 184 GOYFQKLGRCSVRVSVNTANTVLGQPMVTVYRRHGRVYVYIAQVXDYVVTDIQIPVF 243  
Qy 241 TMOKNDNRNSDDEFLKDLPTMFQVLTJHDSHPFLNYSTINVKSPFGDNTGLFVSTNHTVN 300  
Db 244 TMOKNDNRNSDDEFLKDLPTMFQVLTJHDSHPFLNYSTINVKSPFGDNTGLFVSTNHTVN 303  
Qy 301 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTSLGPAGDNPLLSRIPDENCQIN 360  
Db 304 HTYVLNGTFSNLTVKAAAGPCPPPPPPRPSKPTSLGPAGDNPLLSRIPDENCQIN 363  
Qy 361 RYGHFOATTIVGILEVNIIOQMTDVLMPWPBPSSLIIDFVVTQCGSIPTEVCTIIISDPT 420  
Db 364 RYGHFOATTIVGILEVNIIOQMTDVLMPWPBPSSLIIDFVVTQCGSIPTEVCTIIISDPT 423

Qy 421 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALSTLSVDRDRPAS 480  
Db 424 CEITQNTVCSVDVDEMCLLTVRTFNGSGTYCVNLTGLGDDTSLALSTLSVDRDRPAS 483  
Qy 481 PLRMANALISVGCLAFVTVVISLLVYKHKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 540  
Db 484 PLRMANALISVGCLAFVTVVISLLVYKHKHKEYNPIENSPGNVVRSGLSVFLNRAKAVF 543  
Qy 541 FPGNQEKDPLLNQEPKGV 560  
Db 544 FPGNQEKDPLLNQEPKGV 563

RESULT 3  
US-11-090-439-44  
; Sequence 44, Application US/11090439  
; Publication No. US20050266442A1  
; GENERAL INFORMATION:  
; APPLICANT: Squillace, Rachel  
; APPLICANT: Weiner, Michael P.  
; TITLE OF INVENTION: Immortalized Human Tuberous Sclerosis Null  
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof  
; FILE REFERENCE: 24318-502  
; CURRENT APPLICATION NUMBER: US/11/090,439  
; CURRENT FILING DATE: 2005-03-25  
; PRIOR APPLICATION NUMBER: 60/556,344  
; PRIOR FILING DATE: 2004-03-25  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 44  
; LENGTH: 206  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-11-090-439-44

Query Match 34.0%; Score 1029; DB 7; Length 206;  
Best Local Similarity 98.4%; Pred. No. 2.1e-82;  
Matches 182; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGWSSDENDWNEKLYP 60  
Db 1 MECLYYFLGFLLLAARLPDAAKRFHDVLGNRPSAYMRHNLQNGWSSDENDWNEKLYP 60  
Qy 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIIPRCOKEDANGNIYVEKNC 120  
Db 61 VKRGDMRWKNSWKGGRVQAVLTSDSPALVGSNTITFAVNLIIPRCOKEDANGNIYVEKNC 120  
Qy 121 RNEAGLSADPVYVNTWTAWSESDGNGTGSQSHNVFPDGPFPHPGRRWNFIYVPHTL 180  
Db 121 RNEAGLSADPVYVNTWTAWSESDGNGTGSQSHNVFPDGPFPHPGRRWNFIYVPHTL 180  
Qy 181 GOYFQ 185  
Db 181 GLLQ 185

RESULT 4  
US-11-090-439-9  
; Sequence 9, Application US/11090439  
; Publication No. US20050266442A1  
; GENERAL INFORMATION:  
; APPLICANT: Squillace, Rachel  
; APPLICANT: Weiner, Michael P.  
; TITLE OF INVENTION: Immortalized Human Tuberous Sclerosis Null  
; TITLE OF INVENTION: Angiomyolipoma Cell and Method of Use Thereof  
; FILE REFERENCE: 24318-502  
; CURRENT APPLICATION NUMBER: US/11/090,439  
; CURRENT FILING DATE: 2005-03-25  
; PRIOR APPLICATION NUMBER: 60/556,344  
; PRIOR FILING DATE: 2004-03-25  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1





```
RESULT 7
US-10-995-561-818
; Sequence 818, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 818
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-818

Query Match      3.5%; Score 105.5; DB 6; Length 427;
Best Local Similarity 32.4%; Pred. No. 0.072;
Matches 36; Conservative 14; Mismatches 38; Indels 23; Gaps 6;

Qy 317 AAAPGCPPPPPPPRPSKPT--PSLGPAGDNPLSLRIPDENCQINRYGH-----FOAT 368
Db 161 AVAPEPCQPLRSPSLDNPTFPNLGPS-ENPLKRLVPGGEWFEVTAFYRGQVFOQT 219

Qy 369 ITIVGILEVNIQMTDVLMP-----VWPBESSLID-----FVTCQG 406
Db 220 ISCEGLRLVG-SEVGDRTLGWPVTLDPDQMSLTDGVMYSYVRHVLSCLG 269

RESULT 8
US-10-995-561-817
; Sequence 817, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 817
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-817

Query Match      3.5%; Score 105.5; DB 6; Length 452;
Best Local Similarity 32.4%; Pred. No. 0.078;
Matches 36; Conservative 14; Mismatches 38; Indels 23; Gaps 6;

Qy 317 AAAPGCPPPPPPPRPSKPT--PSLGPAGDNPLSLRIPDENCQINRYGH-----FOAT 368
Db 161 AVAPEPCQPLRSPSLDNPTFPNLGPS-ENPLKRLVPGGEWFEVTAFYRGQVFOQT 219

Qy 369 ITIVGILEVNIQMTDVLMP-----VWPBESSLID-----FVTCQG 406
Db 220 ISCEGLRLVG-SEVGDRTLGWPVTLDPDQMSLTDGVMYSYVRHVLSCLG 269

RESULT 9
US-11-192-967-2
; Sequence 2, Application US/11192967
; Publication No. US20050262587A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
```

```
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; TITLE OF INVENTION: Lepidopteran Insects and Methods of Use
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/192,967
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-11-192-967-2

Query Match      3.2%; Score 97; DB 7; Length 1717;
Best Local Similarity 21.2%; Pred. No. 2.5;
Matches 91; Conservative 67; Mismatches 156; Indels 116; Gaps 26;

Qy 154 NVPPDGKPPPH-----PGWRMMFIY--VFHTLQY-----FQKLGRCVRSVNTANVTL 203
Db 434 DVFPFGAASAFYIAPGSGYQRTFIMGTINHTMLDYEDVIFQ-----NIIKVKAVDMNN 488

Qy 204 GPQLMEVTVYRRHGRAVPIAQKDVVV--TQIPIVFTVMFQKNDNRSSDETLKDLPI 261
Db 489 ASHVGEALVY-----VNLIWNDELPIFE---ESSYASAFKETVGAGFPV 530

Qy 262 -----MFDVLHDPHFLNYSYTYKWSFDNTG-LFVSTN-----HTVNHTYVL- 305
Db 531 ATVLALDRDIDVVVHS-----LGNADVILF-IDESTGEIFVSMDDAFYHRQNTL 585

Qy 306 -----NGTFSMLTVKAAAPGCPPPPPRPSKPTPSLGPAGDNPLSLR----- 351
Db 586 RADDTLGDGPHNTVTTLQVIELEDVNNTPPTLRLPR---STPSVEENVPEGEYSREITA 642

Qy 352 -IPDENCQINRYGHFOATITIVGILEVNIQMTDVLMPVWPBESSLIDFVTCQGSPT 410
Db 643 TDPSTLAYLWFEDWDSTWATKQG-RETN-----PTEYVCIVETIITYPTENRGS 692

Qy 411 EVCTIISDPTCEITQNTVCSFVDVD--EMCLLTVRTFNGSGTYCVNLTGLDDTSLA-LT 467
Db 693 AIGRLVVQ---EIRDNT---IDFEFEMLYLTVR-----VRDLNTVIGDDYDEATFT 739

Qy 468 STLISVDPDRDPASPLRWANSALISV-----GCL--AIFVTVISLLVYKKHK-EYNP 515
Db 740 ITIIDMNDNAPI-----FANGTLTQTMVRVLAASGTLIGSVLATDIDGPLYNQVRYTIQ 795

Qy 516 IENSPGNVVR 525
Db 796 RNNTPEGLVK 805

RESULT 10
US-11-193-715-2
; Sequence 2, Application US/11193715
; Publication No. US20050260675A1
; GENERAL INFORMATION:
; APPLICANT: Flannagan, Ronald D.
; APPLICANT: Mathis, John P.
; APPLICANT: Meyer, Terry E.
; TITLE OF INVENTION: Novel Bt Toxin Receptors From
; TITLE OF INVENTION: Lepidopteran Insects and Methods of Use
; FILE REFERENCE: 35718/204664
; CURRENT APPLICATION NUMBER: US/11/193,715
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: US/09/715,909
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/166,285
; PRIOR FILING DATE: 1999-11-18
```

```
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-11-193-715-2

Query Match      3.2%; Score 97; DB 7; Length 1717;
Best Local Similarity 21.2%; Pred. No. 2.5;
Matches 91; Conservative 67; Mismatches 156; Indels 116; Gaps 26;

QY 154 NVFPGKPPPH-----PGRWRNFIY--VFHTLQY-----FQKLGRCVRSVNTANVTL 203
DB 434 DVFFPGAASAFIAPGSGYQRTFIMGTHNTMLDYEDVIFQ-----NIIIKVKAVMNN 488
QY 204 GPQLMEVTVYRRHGRAYVPIAQVKDYYVV--TDQIPVFVTFMQKNDNRSSDETFLKDLPI 261
DB 489 ASHVGEALVY-----VNLINWDELPIFE---ESSYSASFKEITVCAGFPV 530
QY 262 -----MFVLIHDSHFLNYSTINYKWSFGDNTG-LFVSTN-----HTVNHYVL- 305
DB 531 ATVALDRDIDVVVHS-----LMGNAVLYLF-IDESTGEIFVSMDDADFYHRQNTLFVQV 585
QY 306 -----NCTFSLNLTVRAAAGPCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 351
DB 586 RADDTLGDPGPHNTVTTLQVLELDVNNPTTLRLPR-----STPSEVENVPBGYEISREITA 642
QY 352 -IPDENCQINRYGHFOATITIVEGILEVNIIMQTDVLMVPWPRESSLIDFVVTCGSIPT 410
DB 643 TDPDTSAVLNPEIDWDSWATKQG-REYN-----PTEVGCIVIETIYPTENRGS 692
QY 411 EVCTIISDPTCHITONTVCSPVDVD--EMCLTVRRTFNGSGTYCVNLTLGDDTSLA-LT 467
DB 693 AIGRLVWQ--EIRDNV---IDPEFEFMYLTVR-----VRDLNTVIGDDYDEATFT 739
QY 468 STLISVPDRDPASPLRMANSALISV-----CCL--AIPVTVISLIVYKHKH-EYNP 515
DB 740 ITIIDMNDNAPT-----FANGTLTQTMVRVRELAASGTGLIGSVLATDIDGLYNNQVRYTIQ 795
QY 516 IENSPGNVVR 525
DB 796 RNNTEGLVK 805

RESULT 11
US-11-080-991-64
; Sequence 64, Application US/11080991
; Publication No. US20050266437A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Petter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/11/080,991
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: US/10/176,847
; PRIOR FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 583
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-080-991-64

Query Match      3.2%; Score 96.5; DB 7; Length 583;
Best Local Similarity 19.4%; Pred. No. 0.66;
Matches 80; Conservative 53; Mismatches 137; Indels 143; Gaps 16;

QY 205 POLMEVTVYRRHGRAYVPIAQ-VKDYYVVVDQIPVFVTFMQKNDNRSSDETFLKDL----- 259

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-11-193-715-2

Query Match      3.1%; Score 94.5; DB 7; Length 557;
Best Local Similarity 23.1%; Pred. No. 0.92;
Matches 74; Conservative 40; Mismatches 135; Indels 71; Gaps 19;

QY 86 SPALVGSNITFAVNLIPRCOKEDANGNIVYEKNCNEAGLSADPYV--YNTAWSESDG 144
DB 114 APELLGGSVF-----LFPKPKDTLMISRTPEVTCVVVDVSHEDDEVKFNVTV----- 162
QY 145 ENGTCQSHHNVFPDGKPPPHHPGWRWNFIYVFH---TLGQYFQKLGRCVRSVSYNTANV 201
DB 163 ---DGEVHNNAKTPREBQYNSYRVVSVLTVLHQDWLNGKEY----KCKV-----SNK 209
QY 202 TLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTDQI---PVFTMFQKN-----DR 248
DB 210 AL-PAPIEKTISKAKGQPRRP--QVYTLPPSRDELTKNQVSLTCLIVKGFPYPSDIAVWES 266
QY 249 NSSDETFLKDLPIMFVDLIHDPSPHLYNYSINY---KWSFGDNTGLF-VSTNHTVNTVTV 304
DB 267 NQGPENNYKTTFP-----PVLDSGGSFFL--YSKUTVDKSRWQQGN---VFSCSVNHEALHNY 319
QY 305 LNGTSPSLNLTVKAAPG-----PCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 358
DB 320 TQKSLSL-----SPGDKTKHTCPSPAPPELLGGPSVFLPFPKPKDTLMISRTPEVTCV 372

; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1717
; TYPE: PRT
; ORGANISM: Ostrinia nubilalis
US-11-193-715-2

Query Match      3.1%; Score 94.5; DB 7; Length 557;
Best Local Similarity 23.1%; Pred. No. 0.92;
Matches 74; Conservative 40; Mismatches 135; Indels 71; Gaps 19;

QY 86 SPALVGSNITFAVNLIPRCOKEDANGNIVYEKNCNEAGLSADPYV--YNTAWSESDG 144
DB 114 APELLGGSVF-----LFPKPKDTLMISRTPEVTCVVVDVSHEDDEVKFNVTV----- 162
QY 145 ENGTCQSHHNVFPDGKPPPHHPGWRWNFIYVFH---TLGQYFQKLGRCVRSVSYNTANV 201
DB 163 ---DGEVHNNAKTPREBQYNSYRVVSVLTVLHQDWLNGKEY----KCKV-----SNK 209
QY 202 TLGPQLMEVTVYRRHGRAYVPIAQVKDYYVVTDQI---PVFTMFQKN-----DR 248
DB 210 AL-PAPIEKTISKAKGQPRRP--QVYTLPPSRDELTKNQVSLTCLIVKGFPYPSDIAVWES 266
QY 249 NSSDETFLKDLPIMFVDLIHDPSPHLYNYSINY---KWSFGDNTGLF-VSTNHTVNTVTV 304
DB 267 NQGPENNYKTTFP-----PVLDSGGSFFL--YSKUTVDKSRWQQGN---VFSCSVNHEALHNY 319
QY 305 LNGTSPSLNLTVKAAPG-----PCPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 358
DB 320 TQKSLSL-----SPGDKTKHTCPSPAPPELLGGPSVFLPFPKPKDTLMISRTPEVTCV 372
```



Db 210 AL-PAPIEKTISKAKQPREP--QVYTLPPSRDELTKNOVSLTCLVKGFYPSDIAVEWES 266  
QY 249 NSSDETFLKOLFIMFDVLIHDPHFNYSTINY---KWSFGDNTGLP-VSTNHTVNNHTYV 304  
Db 267 NGQPENNYKTPE---PVLDSGSGFFL-YSKLTVDKSRWQGN---VFSCSVMHREALHNY 319  
QY 305 LNGTFSNLTVKAAAPG--PCPPP-----PPPRPSKPTPSLGPAGDNPELSR 351  
Db 320 TOKSLSL-----SPGKPPCPAPELLGGPSVFLFPKPK-----DTLMISR 359  
QY 352 IPDENCQINRYGHFOATITI---VEGI 375  
Db 360 TPEVTCVVVDVSHEDPEVKFNWYVDGV 386

Search completed: December 24, 2005, 04:50:19  
Job time : 20 secs

**THIS PAGE BLANK (USPTO)**